UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PHONOLOGICAL ISSUES IN THE PRODUCTION OF PROSODY BY FRANCOPHONE AND SINOPHONE LEARNERS OF ENGLISH AS A SECOND LANGUAGE

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN LINGUISTIQUE

PAR
MARIE PLOQUIN

FÉVRIER 2009

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PHONOLOGICAL ISSUES IN THE PRODUCTION OF PROSODY BY FRANCOPHONE AND SINOPHONE LEARNERS OF ENGLISH AS A SECOND LANGUAGE

THESIS

PRESENTED

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN LINGUISTICS

BY

MARIE PLOQUIN

FEBRUARY 2009

To Roney,

ACKNOWLEDGEMENTS

First and foremost, I express my utmost gratitude to Henrietta Jonas-Cedergren, my advisor, for her availability, thoroughness, enthusiasm, encouragement as much as her kindness and good humour. Her perceptive comments inspired me, stimulated my curiosity and sharpened my sense of research.

I am also thankful to Claire Gélinas-Chebat and Robert A. Papen for devoting the time to review my work and providing insightful comments. I am also indebted to Hua Lin who kindly agreed to act as the external examiner.

I express my warmest thanks to my husband, Roney Audet, whose sacrifices, compromises and proofreading made the completion of this thesis possible. Furthermore, his encouragement and good sense humour facilitated the whole process.

This thesis is the culmination of years of academic training in which the following University of London, RHBNC, professors contributed most: Dr Mike Routledge who believed in me and made me believe in myself, Dr Marie Landick who instilled in me a taste for linguistics and the sense of rigour and Dr Stephen Gundle whose pedagogical skills inspired me to join the teaching profession. I also thank the linguistics professors at University of Cambridge, Cambridge, UK and at *Université du Québec À Montréal* (UQAM), Montréal, Canada.

The experiments of this thesis were made possible by the participation of the following individuals, whom I give my thanks to: all the subjects of the experiments and, in alphabetical order, Kathleen Brannen, Roisin Dewart, Josephine Eriavaz, Li Lei, Donna Lisney, Heather Newel, Jane Nicholls, Liya Pan, Emily Rosales and

Sydney Wise. A special thank you goes to Vincent Arnaud for his precious help with Praat.

Finally, I thank the following people for their encouragement and understanding:
My colleagues Jaime Demperio, François Desaulniers, Roisin Dewart, Martyna Kozlowska, Carey Nelson, Emily Rosales and Gerald Rosenau.
My parents, Yves and Pierrette Ploquin.
My friends, among them: June Léger, Eleonor Welby, Marie-Anne and Micheline Dion.

TABLE OF CONTENTS

CHAPTER III

PHONOLOGY: RESEARCH & APPLICABILITY

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

[1] The legend of Shen's tunes has been amended here as the reference to the lines in her legend was incorrect with respect to her graph.

CHAPTER 5

LIST OF TABLES

CHAPTER 3

CHAPTER 4

LIST OF ABBREVIATIONS AND ACRONYMS

AM: Autosegmental-Metrical

AP: Accentual phrase

EAL: English as an additional language

EFL: English as a Foreign Language

ESL: English as a second language (in this thesis, encompasses ESL and EFL)

ESOL: English for speakers of other languages

F0: Fundamental frequency

FF: French of France

ip: Intermediate phrase

IP: Intonational phrase

L1: First language

L2: Second language

NAE: North American English

NP: Noun phrase

OT: Optimality Theory

PP: Prosodic phrase

PW: Prosodic word

QF: Quebec French

SC: Standard Chinese

ToBI: TOnes and Break Indices

WFC: Well-formedness condition

ABSTRACT

Accented non-native speech can lead to a lack of comprehension or to the perception of various degrees of foreignness. Prosody, which is now recognized as an important element of accented speech, has been relatively unexplored in second language studies. This contrasts with the growing interest of research on first language prosody.

In this thesis, phonological research is evaluated for its significance to research in second language prosody. Two aspects of phonological theory are investigated: typology and phonological organization. This choice was driven by the belief that prosodic foreignness is due to one of two factors: a difference of typological class of L1 and L2 and a transfer of L1 prosodic features.

The review of research on phonological typology led to the conclusion that, at this stage, no prosodic classification model can be applied to L2 acquisition. More specifically, the investigation shows that some typologies, particularly Pike's stress-time and syllable-time theory, should be dismissed as they hinder progress in research in second language acquisition of prosody.

The second aspect of phonological theory investigated in this thesis is phonological organization. The premise is that differences at the underlying prosodic organization level rather than at the surface feature level are transferred from L1 to L2. The thorough analyses of North American English, French and Standard Chinese yield important organizational differences between North American English and the other two languages. Some of these differences are tested with four experiments.

English prosody by native speakers of French is analyzed in rhythmically simple sentences and in rhythmically more complex sentences. The results show that lexical stress is less of an issue than supra-lexical prosodic stress. Specifically, early and late AP (accentual phrase) rises are shown to be erroneously transferred from French to English prosody. However, the study also shows that, while this error is noticed by native speakers of English, it does not affect their perception of stress placement.

English prosody, as produced by native speakers of Chinese, is analyzed in terms of tone transfer and peak alignment. The results provide strong evidence that speakers of Chinese use Chinese tones when producing English pitch accents; in particular, for the vast majority of subjects, tone 2 (the rising tone) is implemented when producing English rising pitch accents. The final experiment reveals that Chinese

speakers tend to be stricter in the alignment of pitch accents with the corresponding stressed syllables than North American English speakers.

The results of this thesis provide insights into L2 prosodic competence progression and native speakers' perception of L2 prosody. The findings have implications in pedagogical content and format of pronunciation training.

Keywords: phonology, phonetics, prosodic phonology, prosody, rhythm, ESL, Quebec French, French of France, Chinese.

RÉSUMÉ

Un accent de non-natif peut mener à une incompréhension ou à la perception de degrés différents d'accent d'étrangeté[2]. La prosodie, qui est maintenant reconnue comme un élément important de l'impression d'étrangeté, est relativement peu abordée en recherche en acquisition des langues étrangères. Ceci contraste avec l'intérêt grandissant envers la prosodie en tant qu'élément de la langue maternelle.

Dans cette thèse, la recherche phonologique est évaluée quant à sa pertinence dans la recherche sur la prosodie des langues étrangères. Deux aspects de la théorie phonologique sont étudiés : la typologie et l'organisation phonologique. Ce choix est justifié par la présomption générale que l'étrangeté prosodique est créée soit par une différence de typologie entre langue maternelle (L1) et langue étrangère (L2) soit par un transfert de traits prosodiques de la L1.

La critique de la recherche en typologie phonologique conclut que, à ce stade, aucun modèle de classification prosodique n'est applicable à l'acquisition d'une L2. En particulier, l'étude démontre que certaines typologies, en particulier la théorie de l'isochronie accentuelle/l'isochronie syllabique de Pike, devraient être exclues parce qu'elles entravent les progrès en recherche sur l'acquisition et la production de la prosodie des langues étrangères.

Le second aspect de la théorie phonologique étudié dans cette thèse est l'organisation phonologique. La prémisse est que les différences sous-jacentes à l'organisation prosodique plutôt que les différences phonologiques de surface sont transférées de L1 à L2. Les analyses approfondies de l'anglais nord américain, le français et le chinois standard révèlent d'importantes différences phonologiques entre l'anglais nord américain et les deux autres langues. Quatre expériences évaluent certaines de ces différences.

La prosodie de l'anglais produite par des locuteurs natifs du français est analysée dans des phrases rythmiquement simples et des phrases rythmiquement plus complexes. Les résultats démontrent que l'accentuation lexicale est moins problématique que l'accentuation prosodique supra-lexicale. En particulier, il est démontré que les montées de fréquence fondamentale (F0) de début et de fin de syntagme accentuel (SA), typiques du français, sont source d'erreur dans la prosodie de l'anglais langue seconde. Il est cependant montré que cette erreur, bien que remarquée par les locuteurs natifs de l'anglais, n'affecte pas la perception de placement d'accentuation par ces derniers.

---

[2] Le terme « étrangeté » est utilisé dans le sens du mot anglais « *foreignness* », c'est-à-dire avec la signification de « caractère étranger » et non de « caractère étrange ».

La prosodie de l'anglais produite par des locuteurs natifs du chinois est analysée en termes de transfert de ton et d'alignement de pic de F0. Les résultats indiquent que les locuteurs du chinois utilisent les tons chinois quand ils produisent des tons accentuels de l'anglais; plus spécifiquement, la majorité des locuteurs utilisent le ton 2 (ton montant) quand ils produisent un ton accentuel montant. La dernière expérience révèle que les locuteurs natifs du chinois alignent le ton accentuel avec la syllabe accentuée à laquelle elle correspond de manière plus stricte que les locuteurs natifs de l'anglais nord américain le font.

Les résultats de cette thèse génèrent un aperçu de la progression de la performance de la prosodie d'une langue étrangère. Les conclusions comportent des implications sur le contenu pédagogique et le format de la l'enseignement de la prononciation.

INTRODUCTION

The aim of the present study is to bridge research on phonological investigation with second language research. More specifically, this work presents an investigation of prosody conducted in such a way that its results can serve the needs of teachers who deal with the acquisition of prosody by learners of English as a second language (henceforth ESL[1]).

In recent years, many foreign language programs have added pronunciation training to their curriculum. This subject of study was long neglected, partially because foreign languages were not really taught as communication tools but rather as mental exercise material. Another explanation for the lack of pronunciation tuition is that its validity was disputed by some who felt that only extended exposure to the language would help. With increased international travel and exchange, students' wishes and needs to improve their oral skills had to be addressed. English, a very popular second language whose pronunciation – for instance, in terms of syllable structure, quantity of phonemes and rhythm rules – is particularly complex, was one of the first to benefit from this change. Later came the realization that students of English as a second language or as a foreign language remained difficult to understand even when they had mastered the phonemic inventory of English. The natural conclusion was thus that effective pronunciation teaching required the teaching of suprasegmental elements (i.e. prosodic elements such as intonation, rhythm, and stress) rather than uniquely that of segmental aspects (i.e. phonemes and phones).

[1] Throughout this work, ESL is used as a generalized term and encompasses English as a second language as well as English as a Foreign Language. The meaning of ESL in this work is thus equivalent to the meaning of EAL (English as an additional language) and ESOL (English for speakers of other languages) in other works.

This awareness is reflected by the number of ESL course books that either include lessons on rhythm (Lane, 1993, 2005; Dauer 1993) or adopt a "rhythm approach" (Miller, 2000). Nevertheless, the claim to include rhythm is often greater than the actual attention given it. Most puzzling is the approach used to teach rhythm: while most authors realize that they must bear in mind the phonemic/phonetic inventory of the students L1 when teaching segmentals, none seem to consider the suprasegmental elements of their students' L1. For instance, numerous exercises on vowels and consonants in Dauer (1993) are labelled for speakers of particular languages. None of these labels are provided in any of the stress and rhythm sections, which would suggest that speakers of all languages face the same prosodic issues.

Yet, L1 to L2 prosodic transfer is acknowledged:

> Finally, this book is also for language teachers – by comparing the intonational categories and their realizations in the target languages, they can pin down the sources of prosodic interference and transfer. (Jun 2005: 4)

These notions of interference and transfer raise the topics of second language acquisition and of individual language characteristics. Theoretical[2] linguistics research deals with these in a number of ways, notably by describing first language and second language (henceforth respectively L1 and L2) speech. The results of these studies either depict the state of a language or are used to generate a theoretical model. In the case of research into prosody, one can indeed find a number of studies that describe the use of fundamental frequency (F0) or duration or intensity by native or non-native speakers of a language. These correlates are for instance measured by Ueyama (2000) for first and second language English and Japanese, by

[2] "Theoretical linguistics" is here understood, as it commonly is, to mean the branch of linguistics that deals with underlying structures of languages and with the search of linguistic universals. It therefore contrasts with "applied linguistics", the field of study which is concerned with the practical application of findings to solving real-life linguistic issues.

Guilbault (2002) for French produced by native speakers of English and by Hua (2003) for English prosody by Taiwan Mandarin ESL learners.

While these studies provide precious information, they fail to identify the source of discrepancies between L1 and L2 prosodies and consequently do not enable language teachers to draw necessary information to help L2 students.

Gut et al. (2007) point out this lack of exchange between theoreticians and foreign language teachers. The two groups operate in parallel although they have an interest in common: L2 prosody. Clearly, these professionals would mutually benefit from closer cooperation: researchers can provide teachers (via textbook writers) with information on areas to concentrate on and teachers can provide feedback on the application of such information. Other factors impede the exchange between the two professional groups, notably:

> [...] a myriad of competing theories and models dealing with fine-grained details [...] which predict very different acquisition processes and attribute different degrees of importance to particular pedagogical strategies and learner characteristics. Gut et al (2007).

CHAPTER I

STRUCTURE AND DESIGN

1.1. Scope, implications and limitations of this study

This thesis is intended to bridge the gap between L2 prosody research and L2 prosody teaching. The study concentrates on North American English[3] prosody and its production by native speakers of French and native speakers of Chinese. The use of two languages reveals the wish to provide information that extends beyond a very specific linguistic context, information that would thus be likely to be transferable to other languages. This research also removes the difficulty of application in the classroom by steering free of any particular language acquisition model. The information provided can thus be adapted to suit any teaching approach and method[4].

The present study has thus implications not only for the field of linguistics and prosody, but also for the area of second language acquisition and teaching. This investigation offers an in-depth analysis of phonological typology and of prosodic phonology. It also offers a complete account of Chinese prosodic phonology in the Autosegmental-Metrical (henceforth AM) framework which can be put side by side

[3] The term "North American English" is understood as the variety of English spoken in North America, that is to say, "English as it is written and spoken by educated speakers in the United States of America and Canada."(Trudgill and Hannah, 2002:2). Although Canadian English and American English can be recognized as individual varieties of English, the great number of similarities between them, notably in pronunciation and vocabulary, has lead these authors as well as Labov et al. (2006) to group them together.

[4] As is customary in second language teaching, the terms "model" and 'approach' describe a language teaching philosophy that can be applied in various ways in the classroom while the term 'method' describes a set of teaching practices based on a particular theory of language learning.

with those of English by Pierrehumbert (1980) and of French by Jun and Fougeron (2000, 2002).

This work does not provide all the answers with regards to second language acquisition of prosody. Nor does it pretend that the findings will apply to all languages. Within these limitations, this multidisciplinary work aims to map out a path for exchange between professional groups so as to better and further contribute to the advancement of prosodic science and of our understanding of the issues.

1.2. Organization of this study

In chapter 2, different aspects of prosody are described and defined. In view of the various definitions of key terms, the words "prosody", "stress", "rhythm", "intonation" and "tone" are defined. Such aspects as the functions and uses of prosody, the perception of prosodic foreignness and the teaching of prosody are discussed. The final section presents major acquisition models as well as the L1 transfer debate and introduces the concept of tonal transfer.

In chapter 3, phonological research is studied for help in better understanding prosodic similarities and differences between languages. The first investigation explores the typology of rhythm, which aims to categorize languages on the basis of their rhythmic similarity. The principle of this typology is that the more the rhythmic similarities between a given L1 and English, the easier the acquisition of English rhythm. All major typologies are presented and assessed from a general linguistic perspective but also in the context of the acquisition of English pronunciation by ESL students.

Second, the phonological organization of rhythm for the three languages (English, French and Chinese) is described. The premise is that phonological features of prosody might be transferred from L1 to L2. After a thorough description within

the AM framework, the three prosodic phonologies are compared to highlight areas that would be favourable for prosodic transfer from French to English and from Chinese to English.

Chapter 4 presents four experimental studies conducted to verify the hypotheses of transfer posited in the previous chapter. Two experiments are carried out on French ESL speakers: one to evaluate their production of rhythmically simple sentences and one for rhythmically more complex English sentences. The other two experiments are carried out on Chinese ESL speakers: one to evaluate the production of Chinese tones in place of English pitch accents and one to analyze pitch alignment.

This dissertation concludes in a fifth chapter which attempts to integrate the results of the experiments in relation to the theoretical findings. The focus of the conclusion is placed on the possibility of integration and exploitation of the findings in the ESL classroom.

CHAPTER II

PROSODY

2.1. Prosody: Definition and description

Prosody is a general property of language and the term "prosody" refers to the rhythmic aspect of language. Prosodic features are not permanently associated with particular segments and are thus suprasegmental (i.e. above the segment).

The perception of prosody is created by the fluctuation of a number of measurable properties, i.e. acoustic attributes:

- Fundamental frequency (F0), which indicates pitch variation;
- Intensity, index of volume and force;
- Length, index of the duration of production of a phoneme;
- Pauses, which are sound free periods between syllables;
- Tempo which indicates rate of speech.

Prosody encompasses stress, rhythm, and intonation. As the use and definitions of these terms vary, the terminology, as used in this study, is here defined:

- *Stress* refers to the relative prominence of a syllable in a word and is sometimes called "word accent". Prominence of a syllable may be achieved by greater or lower frequency, pitch patterns, greater or lower intensity, greater or shorter non-intrinsic segmental[5] length. Stress is thus the quality of a syllable that stands out in comparison to others. In this work, "stress" is thus used to refer to the prominence of a syllable in word (including compound words).

[5] Non-intrinsic segmental length refers to the effect of prosody on segment length.

- *Rhythm* refers to the perception of repetition of events across time; these events may be F0 valleys and peaks (of stressed syllables for instance), intensity rises and falls, non-intrinsic segmental length variation, absence/presence of pauses and length variation of pauses. In this work, rhythm is used to refer to these events across segments such as phrases and clauses.[6]

- *Intonation* refers to pitch movement (i.e. contour) applicable to the entire sentence. This contour is normally continuous but may contain dislocations at the level of phrase and clause boundaries. In this work, intonation is used in this sense, that is to say, as F0 variation in the speaker's voice relevant to the meaning of the utterance as a whole[7].

Tone, which refers to pitch movement applicable to the syllable, is also a suprasegmental feature because it spans more than one segment. It is however not a prosodic feature because it is associated with a lexical item (for example Chinese, Thai and Yoruba) or a grammatical function (in many African languages). Yet, because tones are created by pitch variation, they have an impact on prosody, especially on rhythm.

2.2. Prosody: Functions and uses

One of the functions of prosody is to reflect syntax. Although the assumption of an isomorphic correspondence between syntactic phrases and prosodic phrasing has long been dismissed, prosody may for instance reflect the syntactic segmentation of

[6] Rhythm is thus understood here as what Cutler (1984:82) calls "accent" and Fox (2000:115) refers to as "accentuation": "in the rather broader sense of the overall organization of speech in respect of accents"

[7] This differs considerably from the concept of intonation by Pierrehumbert (1980) for whom the term encompasses all prosodic features (rhythm, stress, pauses and intonation). Pierrehumbert's understanding of "intonation" is therefore equivalent to that of "prosody" here.

a sentence, or the nature of the word. A well-known example of the former is the difference between the two interpretations of the French sentence:

(1)a *La belle ferme le voile*
[Det Adj N] [Pro V]
The nice farm hides it.

(1)b *La belle ferme le voile*
[Det N] [V Det N]
The belle draws the veil

While the written form is ambiguous, the phonetic realization is unequivocal as the noun *ferme* in the first interpretation and *belle* in the second are the most stressed elements of the sentences, as illustrated in figure 2.1.

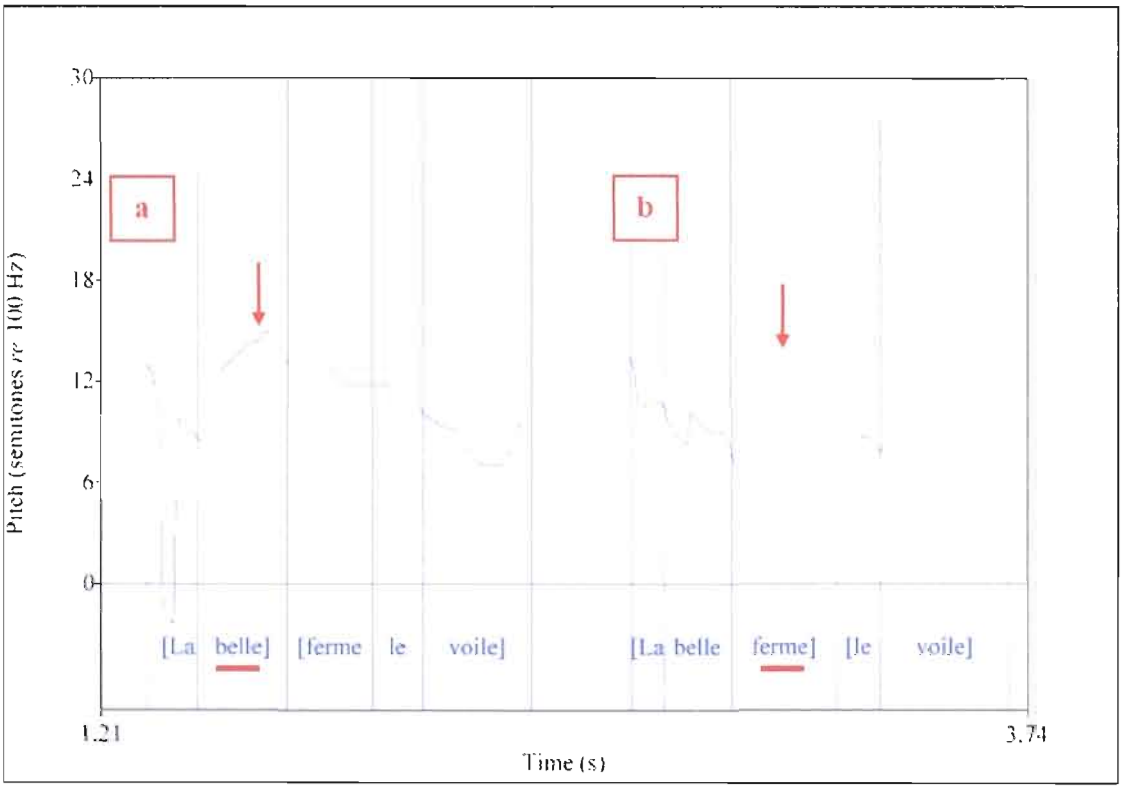


Figure 2.1: Prosodic difference between the realizations of the sentence "*la belle ferme le voile*" according to syntactic segmentation. In (a), the highest F0 value is aligned with the noun *ferme*. In (b), the highest F0 value is aligned with the noun *belle*. (a) *The belle draws the veil; (*b) *The nice farm hides it*

Each language uses a selection of rhythmic attributes to mark syntactic information. Speakers of a given language have little, if any, leeway, in this respect.

Prosody can also provide information about the function of a sentence (interrogative, affirmative, exclamatory or imperative). This is normally conveyed by intonation, that is to say, by F0 movements. Again, speakers have little latitude for personal variety in the sense that, if a rising intonation is required for an interrogative sentence for instance, speakers may only slightly increase or slightly decrease the rise. Changing the contour to a descending one would change the function of the sentence. In English, ascending too quickly or too high would load the question with a sense of incredulity on the part of the speaker.

Indeed, the speaker's attitude and emotional state are also conveyed in prosody (affective prosody). For instance, anger, happiness, ambivalence, and disgust are clearly recognized in speech. The prosodic information is conveyed by intonation and by rhythm, through modulation of all the acoustic cues mentioned earlier as well as speech rate. Fear, for instance, is characterized by an increase of F0, pitch range, and speech rate (Whiteside, 1998).

Prosody also conveys the speaker's communicative intent, for example, intent of derision (through sarcasm) or intent of stimulation (through infant-directed speech, "motherese"). The latter, for instance, is realized by elevated pitch, slower speech rate and swooping contours. Speech modulation indicating affect (i.e. speaker attitude, emotional state, and communicative intent) is unconsciously realized by speakers in the sense that they don't think about it. The term "unconsciously" should not however be associated to "universal", as the prosodic cues relating to each type of affective prosody may vary from language to language, from regional variety to regional variety, or from sociolect to sociolect. Speakers therefore produce affective prosody as unconsciously as they do intonation, after assimilation of their language's use of correlates. The use of rising intonation at the end of a statement, a phenomenon called High-Rising Terminal (HRT), "uptalk", or

"upspeak", is quite common in Canada and the United-States, especially among younger people. This intonation is similar to what is normally recognized as Standard English interrogative intonation or sometimes as a continuation rise. A similar phenomenon is observed in Northern British cities such as Belfast, Liverpool, and Tyneside (Hirst, 2005; Grabe and Post, 2002).

Finally, speakers use prosody to indicate focus, contrast, and emphasis. This is normally realized by F0 contours and could be seen to be more intentional on the part of the speaker.

2.3. Foreignness of second language prosody

With the many variables (uses and functions of prosody and prosodic cues), one could almost suggest that the number of combinations of these variables is very high and that most combinations would be acceptable. If that were the case, prosody would not be judged important in the speech of ESL learners. However, there is no question that native listeners easily perceive if their interlocutor is a native speaker of English with a regional accent or a speaker of English as a foreign or second language (EFL or ESL). This indicates that speakers have a clear notion of what is acceptable in their language and what is not. Of course, all departures from the prosodic norm are not equal. Indeed, from the description of prosody offered above, one can imagine that deviations affecting rhythmic information relating to syntax and sentence segmentation will impair intelligibility. Other prosodic discrepancies might give an incorrect impression about the speaker's attitude and intention. The latter, incorrect affective prosody, should not be considered less important than impaired intelligibility, especially in the context of immigrants' social and professional integration.

A few authors (Anderson-Hsieh et al. 1992, Munro 1995, Kamiyama 2004) have considered the importance of suprasegmental features in the perception of foreignness. Boula de Mareuil and Vieru-Dimulescu (2006), for instance, have found that prosody is as important as (in one experiment) or more important than (in another experiment) segmental information in the detection of foreign accent. A major issue is to understand which prosodic divergences make non-native speakers sound foreign. To write a complete inventory of such differences would certainly be utopian. Notwithstanding, one of the objectives of this study is to identify some of these prosodic divergences.

2.4. Teaching of ESL prosody

Second language programs do not offer courses specifically designed to teach prosody. As is explained below, when prosody is taught, it is as part of pronunciation classes. For this reason, the following discussion centers mainly on pronunciation although it reverts to prosody when appropriate.

2.4.1. Students' perspective

Often times, learners of EFL consider certain linguistic components more essential than others in enabling them to gain proficiency in the language. For instance, for most, the acquisition of grammar and vocabulary is inescapable while that of other elements such as pronunciation might not be seen as warranted, at an early stage at least. Pronunciation is sometimes seen as a skill that can be overlooked and that should be acquired only by those who wish to refine their mastery of the language. In other words, it is seen as some form of gilding. One of the reasons for this is that, in many languages, pronunciation is closely related to orthography. Students may not understand at the early stages of their learning that the relation between English spelling and pronunciation is far from straightforward and that not mastering

pronunciation prevents them from being understood rather than just make them sound foreign.

Furthermore, many seem to think that the likelihood of acquiring pronunciation skills is quite minimal because of the general belief that pronunciation is not something that one learns by means of lessons but through exposure to and practice of the language. Thus, pronunciation is also often seen as a hindering but uncontrollable part of the language.

Furthermore, students largely associate pronunciation with phonemes. Very few would consider prosody an issue. One explanation for this is that students are often unaware of the prosody of their first language, let alone, their prosodic L2 errors.

> The notion of (lexical) stress is indeed very elusive for French natives. They only discover the existence of that unnatural and unnecessary complication when they have to learn a foreign language. (Vaissière, 2002:5)

2.4.2. Professionals' perspective

EFL students however are not the only ones to question the merits and validity of teaching/learning pronunciation. In his summary of the debate over the benefits of teaching pronunciation, Otlowski (1998) shows the contrast between points of view. Indeed, Suter and Purcell (1980:286) argue that their research shows that "the attainment of accurate pronunciation in a second language is a matter substantially beyond the control of educators". Pennington (1989) on the other hand believes that Suter and Purcell's findings are possibly due to the teaching approach and/or quality but that teaching pronunciation is quite feasible and valuable. Between these conflicting opinions, a number of authors totally evade the issue of qualifying the value of the teaching of pronunciation. Stern (1992:112), for instance, simply mentions that 'there is no convincing empirical evidence which could help us sort out the various positions on the merits of pronunciation training.'

2.4.3. Teaching quality

The issue of teaching quality is certainly worth considering. Indeed, in many instances, the students' first EFL experience takes place in their (non-anglophone) country where teachers do not master a standard pronunciation of English – be that standard pronunciation North American, English, Australian, or other. Clearly though, whether these teachers choose to not teach pronunciation because they feel unable to or to teach it despite their own deficiency, the results with regards to student progress are likely to be equally poor. Of course, the simple fact of being a native speaker of English does not guarantee adequate competence for teaching the segmental and suprasegmental features of pronunciation either. Indeed, like many other domains of a language, pronunciation requires specific knowledge and training on the part of the instructor. With respect to prosody, the lack of availability of training material in textbooks to assist instructors further accentuates the problem.

2.4.4. Prosody in textbooks

From the observation that English ESL students remained difficult to understand even when they mastered the phonemic inventory of English came the realization that pronunciation teaching should include the teaching of prosody rather than uniquely that of segmental aspects (i.e. phonemes and phones).

This awareness is reflected by the number of ESL textbooks that are advertised as either including lessons on rhythm (Lane, 1993, 2005; Dauer 1993) or adopting a "rhythm approach" (Miller, 2000). The claim is however often stronger than the actual attention given to suprasegmental features. Most perplexing still, is the approach used to teach rhythm: while most authors realize that they must bear in mind the phonemic/phonetic inventory of the students L1 when teaching segmental features, none seem to consider the suprasegmental elements of their students' L1 when teaching prosody. For instance, numerous exercises on vowels and consonants

in Dauer (1993) are labelled for speakers of particular languages. None of these labels are provided in any of the stress and rhythm chapters.

Furthermore, a number of textbooks use nursery rhymes or poetry to demonstrate rhythm when, in fact, structured rhythm is what differentiates verses from prose. (See chapter 5 "general discussion" for further discussion).

2.5. Acquisition of L2 prosody

The final section of this chapter presents major acquisition models as well as the L1 transfer debate and introduces the concept of tonal transfer.

2.5.1. Review of major acquisition models

The fact that speakers produce phonological features of their first language in their second language is undisputable. This phenomenon is called transfer and, at the phonological level, materializes as what native speakers[8] of that second language perceive as a foreign accent. As specified by Gass and Selinker (2008), it is impossible to not consider the speakers L1 in second language acquisition of phonology: "most work in which transfer was minimized recognized the inevitability of using NL [native language] in the area of phonology" (Gass and Selinker, 2008:181)

Lado (1957) was among the first to propose to look at these differences and similarities between L1 and L2 in order to anticipate foreign language learners' difficulties and the level of these difficulties.

[8] "Native speaker" here refers native speakers as such but also to non-native speakers proficient enough to recognize native phonological features from non-native ones.

> [...] *the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult* (1957:2)

This is the basis of the Contrastive Analysis Hypothesis (CAH) in which easy acquisition is known as "positive transfer" and difficult acquisition, "negative transfer". It must be understood here that "easy acquisition" is assumed to lead to "successful acquisition" – and in turn to positive transfer – and "difficult acquisition" to "unsuccessful acquisition" and negative transfer. Taking support on this model, and from the belief that language is a set of habits (i.e. behaviours), the Audio-Lingual Method was created. When linguists such as Chomsky discredited behaviourism, by association, Contrastive Analysis also lost credit. Furthermore, the predictions of what has come to be known as the "strong version" of CAH were untenable. In particular, it became clear that CAH could not predict all difficulties and that some predicted errors did not occur. However, L1 to L2 phonological transfer is undeniable because identifiable even by the layman.

> *It is for this reason that the Contrastive Analysis Hypothesis in phonology was not abandoned with the same rigor as in syntax. Rather, the attempt was to reconfigure it and incorporate additional principles.* (Gass and Selinker, 2008:181)

This reconfiguration is what came to be known as the *weak* version of CAH which presents an *a posteriori* view of errors and thus has *explanatory* power as opposed to the *a priori* view of the strong version that claims *predictive* powers. This weaker version, which continues to be used, is therefore less of a hypothesis and more of a methodology: contrastive analysis, also called error analysis. Subsequently, the term "interlanguage" was used, first by Corder (1967) and Selinker (1972), to describe the transitional language of learners acquiring a second language. The characteristics of interlanguage include not only transfers but also features such as overgeneralization and simplification; features that cannot be explained as transfer.

The morpheme order studies that surfaced in the 70s were concerned with establishing the order in which syntactical features were learnt. It appeared that features were acquired in relatively the same sequence by both native speakers and non-native speakers and regardless of the latter's L1. The strong hypothesis originally posited was that there is a "natural order" and that L1 and L2 acquisition processes of the morpheme order studies were equivalent:

> [...] this hypothesis was rejected largely because the basic question of what is being acquired in SLA was limited here to a list of isolated English morphemes, with no principled relation to other aspects of English or to other languages, and also because of weaknesses in the research methodology. (Saville-Troike 2006:44)

Natural order acquisition was also one of the five hypotheses in Krashen's 1981 Monitor Model. This approach was largely adopted in the 80s although the hypotheses and claims were difficult –if not impossible– to prove. The four remaining hypotheses are:

- The acquisition-learning hypothesis, central to the model, which distinguishes acquisition (subconscious) from learning (conscious), that is to say unconscious acquisition through exposure and use of the language from conscious learning about the language through studies for instance;
- the Monitor hypothesis, which follows from the acquisition-learning hypothesis, posits that conscious acquisition only serves as a monitor and that true acquisition comes from using the language in meaningful communication;
- the Input hypothesis, which specifies that acquisition is only possible if the input is one level higher (i.e. more difficult) than the student's competence level;
- the Affective Filter hypothesis, which specifies that language acquisition is positively affected when positive affective or emotional factors (e.g. self-assurance) are in play and negatively affected when negative ones (apprehension) are.

The impact of Chomsky's theories was also felt in second language acquisition. Understandingly, the idea that humans have innate Universal Grammar (UG) knowledge is quite appealing as it would mean that the learner's understanding of some Universal rules could be relied on and that only language specific parameters would have to be learnt. In this sense, positive transfer implies that a parameter setting is the same in L1 and L2 and negative transfer that the setting is different. Interlanguage is then the progressive resetting of L1 parameter settings for L2 settings. This view has been adopted in the approach by White (1989), for instance who concentrates more on the morphology/phonology interface. While the author considers that L1 and L2 and interlanguage grammars are different, she maintains that Universal Grammar is still in play, or at least, that it is useful:

> If we are interested in understanding what it means to attain linguistic competence in second language, then it is crucial to have a theory of linguistic competence to provide a general frame of reference. This is true regardless of the question of whether L2 learners have access to UG; even if it turns out that L2 learners do not have the complex knowledge that would be expected under the UG hypothesis, this is something that can only be determined by investigating universal principles and parameters as they are isolated by linguists (White, 1989:183)

While it is generally accepted that UG alone cannot explain all second language acquisition and processes, it has gained at least some validity from experimental investigations. Coppieters (1987:1) for instance found that

> [...] near-native speakers diverge less from native speakers in formal features, such as those currently covered by studies in Universal Grammar, than in "functional" or "cognitive" aspects of grammar.

Of course, how those speakers acquired the different aspects of grammar might also be responsible for the difference. For instance, greater importance given to the teaching of formal features than that given to functional features would also explain the result. Furthermore, the depth and length of exposure to different types and aspects of grammars are likely to have some impact on the non-tutored student's competence.

The influence of Universal Grammar is such that it is totally embraced by many and is integrated in more recent models. Major (2001) proposed the Ontogeny Phylogeny Model which accounts for the various stages of interlanguage as varying ratios of L1, L2 and UG features. The basic claim of the model is that as the learner's interlanguage develops, the influence of L1 decreases, that of L2 increases and that UG first increases and later decreases. This model embraces the notion of Universal Grammar but also those of similarity and of markedness (explained below). This model is notable for its integration of phonology.

While UG has a strong influence in linguistics, other models, with more functional approaches, have also been proposed. Two of these are presented here.
Eckman's (1977) Marked Differential Hypothesis (MDH) was the combination of typological markedness and CAH. Markedness relates to the level of specificity: the more common the phoneme among the world's languages, the lower its markedness and vice-versa. MDH thus predicts that:

> (a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult;
> (b) The relative degree of difficulty of the areas of difference of target language which are more marked that the native language will correspond to the relative degree of markedness;
> (c) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.
>
> (Eckman 1977: 321)

Flege (1995) fashioned the Speech Learning Model to specify the consequences of difference between L1 and L2 phonemes. Indeed, the author explained that when differences are small, i.e. when two sounds are relatively similar, the learner will not acquire the target language sound easily. On the other hand, when two sounds are very different, the learner will be able to perceive and acquire the L2 sound because the difference is more obvious.

Although Lado did include a small part on stress and rhythm, no specific model of acquisition of prosody has been proposed. Prosodic transfer, in the wider sense of L1 prosody having some affect on the production of L2 prosody, is however acknowledged:

> Finally, this book is also for language teachers – by comparing the intonational categories and their realizations in the target languages, they can pin down the sources of prosodic interference and transfer. (Jun 2005: 4)

2.5.2. Prosodic transfer

The few authors who have studied prosodic transfer have concentrated on comparing correlates of suprasegmental features. Hua (2003), for instance, looked at the use of duration, intensity and pitch in the production of English stress by native speakers of Chinese. Ueyama (2000) investigated the use of different correlates in English and Japanese, both as first and second languages. A few studies also explored the effect of language experience and age of exposure to the L2 on suprasegmental features. Trofimovitch and Baker (2006) for instance found that the former has a positive effect on stress timing while the latter affects speech rate, pause frequency, and pause duration.

2.5.3. Tonal transfer

The research on tonal transfer is even scarcer than that on prosodic transfer. Although, as seen earlier, tones are not prosodic features, they are suprasegmental and thus of importance in L2 prosody acquisition and production. A few experiments have shown, for instance, that native speakers of a tonal language use tones in their perception of non tonal languages. Chen's (2007) study, for instance, suggests that Chinese speakers learning Spanish interpret Spanish stress as a combination of a rising and a falling tone.

2.6. Summary

The above examination of prosody has highlighted how the fluctuation of acoustic attributes such as F0, intensity and length modulates stress, rhythm and intonation to customize our speech in terms of syntactic segmentation, function, affect and semantic prominence. It also discusses the place and quality of prosody in second language teaching and offers a review of the more prevalent acquisition models although none were designed specifically for, and very few deal with, prosody.

In this research, an analysis will be carried out in order to study English prosody produced by native speakers of French and by native speakers of Chinese. That is to say that the realization of English prosody, at their level of competence, will be observed. How they reached this level of competence is thus of no interest here and no effort will be made to prove the adequacy of any acquisition model. The interest lies instead in the origin of the mistakes made, and more specifically, the possible transfer of L1 prosodic phonology features into the L2 (English) prosodic phonology. In this perspective, the phonologies – underlying systems of prosody – rather than prosodic correlates – surface manifestations of those systems – will be studied. This is done in the next chapter in two ways: first with a study of the typologies of rhythm that have been proposed to explain prosodic foreignness and acquisition difficulties. Second, this is achieved by an in-depth description and subsequent comparison of the phonologies of English, French and Chinese.

CHAPTER III

PHONOLOGY: RESEARCH & APPLICABILITY

3.1. Introduction

Second language prosody and prosodic errors are commonly attributed to one of two factors: a difference of typology between the two languages or a transfer of L1 prosodic features. Both aspects are studied in this chapter, from the perspective of linguistic research and with the aim of explaining errors produced by learners of English as a second language (ESL).

3.2. Typology of rhythm

A typology is a system aimed at organizing data into groups according to a specific attribute or a set of attributes. In accordance with this general description, the object of a typology of rhythm is to congregate languages that share the same type of rhythm and to distinguish the different types of rhythm. To be most efficient, a typology needs to have as few groups as possible because a large number of groups would defeat the purpose of typology. In addition, the most effective and indubitable typologies are those that take into consideration the smaller numbers of variable attributes. Indeed, the more the number of attributes considered, the greater the chances of items belonging to two or more groups, which, again, defeats the purpose of typology.

An efficient typology of rhythm should therefore enable the clustering of languages which are rhythmically very similar (if not identical) within the clusters but rhythmically very different across clusters.

3.2.1. Review of typologies of languages

Rhythmic classification has long been associated with the concept of timing and isochrony; indeed, it was originally proposed that a classification based on the length of time elapsed between stressed syllables was possible. In the following paragraphs, the different theories and investigations in this field are described and evaluated with regard to phonetic/phonological language classification and, more specifically, with reference to their usefulness in the Teaching of English as a Second Language (TESL).

3.2.1.1. The stress-timed syllable-time theory

There is a long tradition of rhythmic classification, although none was meant specifically for second language acquisition. The most popular division, first discussed by Pike (1945) with regards to rhythmic units[9] only, and extended to languages by Abercrombie (1967:96-98), is based on a discrimination of languages according to their timing. Timing was selected as the discriminating feature because English seems isochronic, that is to say that it appears to display equal inter-stress intervals. Pike's dichotomous categorization separates *stress-timed languages*, languages in which significant variation in syllable length is observed and in which stressed syllables occur isochronously, and *syllable-timed languages*, languages that display a regular pattern of syllable length and in which stressed syllable occur irregularly as their occurrence depends on the number of syllables in the rhythm unit. Ladefoged (1975) also proposed to add a third category, for *mora-timed languages*, in which rhythm would be based on the recurrence of the mora[10]. Early work in the 60's and 70's seemed to confirm the theory and English, French, and Japanese were accepted as the prototypical stress-timed, syllable-timed, and mora-timed languages, respectively.

---

[9] A rhythmic unit, according to Pike (1945:34) in Couper-Kuhlen (1993:8) is "[a] sentence or part of a sentence spoken with a single rush of syllables uninterrupted by a pause"

[10] The mora is a subdivision of a syllable weight: one mora syllables are light and two morae syllables are heavy.

This classification however does not stand to scrutiny as empirical research has shown that some languages have features of both stress-timed and syllable-timed languages or of neither. Spanish, for instance, is considered to be syllable-timed by Pike (1945), Abercombie (1967), Hockett (1958), Gili Gaya (1940) (in Gutiérrez-Díez, 2001), but stress-timed by Navarro Tomás (1916, 1917, 1921, 1922) (in Gutiérrez-Díez, 2001), while Pointon (1980), using Gili Gaya's data, affirms that Spanish has no regular rhythm whichever event (syllables or stresses) is considered and thus is neither stress- nor syllable-timed (in Bertrán:1999:106).

Despite the generally accepted view that experimental evidence has refuted this stress- and syllable-timing classification, "It remains the prevailing view and still features in accounts of the rhythms of speech because no other hypothesis matches its deceptively bewitching power." (Cauldwell, 2002).

The reason for which the empirically discredited theory is still generally referred to and is even used as a premise in very recent work (see Jun 2005) lies probably less in witchcraft than in the lack of a better hypothesis. Also, as the following quote illustrates, some linguists have an aversion to disclaiming what many reputable linguists have claimed. "Clearly, for so many people to agree, the theory must be saying something true, at least about English." (Dauer, 1983:52)

In view of what has been said here, some linguists have suggested modifications to the classification while others have proposed other rhythmic typologies.

3.2.1.2. The timing continuum

Dauer (1987) proposed that "stress-timed" and "syllable-timed" were not categories but rather extremes of a continuum in which any language would be placed[11]. The

[11] The idea of a more or less stress-timed language was raised, albeit not developed, by Delattre (1965) who said that Spanish was syllable-timed in contrast to English.

stress-timed and syllable-timed dichotomy remains the basis of her classification but the scale warrants rhythmic differences between languages of a same group. Also, with the idea of a scale, these differences can be considered as the consequence of the interaction between phonological, phonetic, segmental, and suprasegmental phenomena which are believed to accentuate or diminish the contrast between stressed and unstressed syllables and therefore make it sound more or less stressed-timed. It must be mentioned that Dauer does away with the notion of syllable-timing as, for her, a low stress-timing score can be related not just to syllable length but also to "patterns of tone, of syllable or vowel length, or even the repetition of certain segmental or grammatical features" (1987:448). Each language is thus evaluated for the status of each of a series of components and is given scores which add up to placing the language on the rhythm scale.

The suggested components of language rhythm are, according to Dauer (1987):

Length

- Duration
  - (+) Stressed syllables are substantially longer than unstressed ones.
  - (0) Stressed syllables are only slightly longer than unstressed syllables.
  - (-) Stressed syllables have the same duration as unstressed syllables.
- Syllable structure
  - (+) Great variety of syllable types and complexity.
  - (-) Limited variety of syllable types (CV and CVC) and presence of phenomena such as cluster simplification, epenthesis, and liaison to simplify complex syllable structures.
- Quantity (if the language has segmental quantity distinction)
  - (+) Quantity distinction possible only in stressed syllables.
  - (0) Quantity distinction usually in stressed syllables but occasionally possible in unstressed syllables.
  - (-) Quantity is not conditioned by stress.

Pitch

- Intonation
  - (+) Intonation related to stressed syllables; emphasis and contrast affect primarily stressed syllables.
  - (-) Intonation is not related to stress; emphasis and contrast affect stressed and unstressed syllables.
- Tone
  - (+) If present, tones on stressed syllables only
  - (0) Tone is neutralized or subject to sandhi in unstressed syllables.
  - (-) Tones on all syllables.

Quality

- Vowel
  - (+) Full vowel sounds in stressed syllables and reduced or centralized vowel sounds in unstressed syllables.
  - (0) Few and not necessarily centralized unstressed vowel sounds
  - (-) Vowel sounds remain the same in both stressed and unstressed syllables. Changes in quality not related to stress
- Consonant quality
  - (+) Consonant sounds precisely articulated in stressed syllables but reduced to special allophones or neutralized in unstressed syllables
  - (-) Consonant sounds remain the same in both stressed and unstressed syllables.

Function of stress

- Nature of stress
  - (+) Lexical stress is "free"; moving the stress can result in a new word.
  - (0) Lexical stress is "fixed"
  - (-) no lexical stress

In order to establish the score of a language, the first status of each feature is given a '+', the middle one (if any) is given a '0', the last one is given a '-'. Some of the assumptions on which this classification is based can be disputed; one of these is that all features have the same importance or weight. It would be possible, for instance, to imagine that some of the features are subcomponents of a more general component. For instance, one could say that vowel quality and length depend on syllable structure and therefore are subcomponents of it.

Another issue with Dauer's proposed classification scale is that two languages could obtain the same score without sharing the status of a single feature.
Finally, if Dauer's scale separates languages in terms of syllable structure and in terms of the presence and consequence of stress, it fails to show stress-timing. Indeed, the author concludes: "The more pluses a language has, the more likely we are to say that the language has "strong stress" ("dynamic" or "expiratory" accent) and is "stress-timed"". (Dauer, 1987: 449)

The problem here is that, from a system based on the presence or absence of stress, she concludes something about timing when the link between stress and timing needs demonstration. Furthermore, she infers from (or supplements) this conclusion that speakers of stress-timed languages are expected to easily identify stressed syllables because they use such syllables to segment the flow of speech: "We would expect that naïve native speakers [...] could fairly consistently identify accented syllables in continuous speech." (Dauer, 1987: 449). This correlation needs demonstration too.

On the other hand, one very positive aspect of Dauer's scale is that it breaks down rhythm into features rather than assumes it to be an indivisible entity. This is particularly relevant in ESL research as it provides a selection of rhythmic features

from which to find those that reveal the differences between English and the students' L1, as is currently done for the segmental features of the language.

3.2.1.3. Phonological and phonetic rhythm

In a 1993 paper, Auer puts forward a new model to answer his question "*Is a Rhythm-based Typology Possible?*" . Auer's model is also based on the stress-syllable-timed model although he clearly states that despite "*a considerable amount of phonetic work, this work has remained inconclusive up to the present day*" (1993:3). For him however, the impossibility of demonstrating the validity of the model lies in two facts: first, Abercrombie's (and predecessors') hypothesis of isochrony being restricted to the phonetic level and second, syllable duration being used as sole criterion in discriminating between stress- and syllable-timing. Auer thus proposes to use a set of parameters which take into account the phonological as well as the phonetic level.[12] Working from the basis that a tendency to keep the duration of the foot/word constant is typical of stress-timed languages and the duration of syllables constant is typical of syllable-timed languages, he establishes the phonological traits expected to be associated with the two types of languages (figure 3.1). It should be noted that this division of phonological traits clearly reflects Dauer's organization of features. Auer's word-rhythm traits are those to which Dauer confers a "+" -which amount to stress-timed features- and Auer's "syllable-rhythm" traits are Dauer's "0" or " - " features.

Auer then puts to the test his phonological word- syllable-rhythm (re)model by investigating the predicted correlations between phonological and phonetic parameters of 34 languages.

---

[12] In order to show the phonological, multi-parametrical approach (rather than a purely phonetic one), Auer adopts Donegan and Stampe's (1983) terms "word-rhythm" and "syllable-rhythm" to replace Pike's "stress-timing" and "syllable-timing" respectively.

Syllable-rhythm	Word-rhythm[13]
No accent-dependent reduction	Reduction of non accented syllables in quality and/or duration
[± long] in consonants and vowels of all syllables possible	No [± long] distinction in non-accented syllables
Tone possible	No tone (or non accented syllables are "neutral")
Simple syllable structure	Complex syllable structure,
Open syllables	sonority scale disobeyed
Few assimilations	Frequent assimilations, dissimilation
Syllable division unambiguous	Syllable division ambiguous and variable
No word-related phonological processes	Word-related phonological processes
External = internal sandhi	External ≠ internal sandhi
Vowel harmony possible	No Vowel harmony
Phonetically weak word accent or none at all	Phonetically strong word accent realized by pitch (and other prosodic features)
Word accent (if any) fixed, no grammatical functions	Word accent assigned by complex rules referring to syllable structure, partly morphologized, or free, may have grammatical functions
Geminates possible	No geminates
No central ("reduced") vowel phonemes	Central vowel phonemes possible

Figure 3.1: *Auer's summary of phonological traits associated with syllable-rhythm and word-rhythm (reproduced from 1993:11)*

The approach to typology which he chooses is holistic (all features are connected), deductive (thus the types are idealizations and of which no concrete realization may exist), phonological and phonetic, inclusive of phonological processes and inventories of sounds, and prototypical (thus, a language can differ –or even move away– from a prototype for diachronic reasons. Although Auer concludes that

[13] See footnote 12.

"these results provide positive evidence for the model proposed" (1993:88) he feels that the model should be further revised to improve its correspondence with the findings, and this, for the following reasons:

- The position of the word accent as the central parameter is not acceptable because the expected correlation between word accent and vowel reduction and/or shell complexity[14] – which derive from word accent – is not supported by the results.
- The assumption of phonetic or phonological duration, central to the model, is not corroborated by the findings, especially with regards to parameters such as tone/vowel harmony or word- vs. syllable-related processes. Auer thus entertains the idea of removing all reference to duration from his model. This leads him to conclude that "the notion of rhythm, which is intrinsically linked to duration [...], would have to be abandoned as well.' (1993:89)
- The three tone languages investigated show that tone assignment is related to phonological words while the word-rhythm syllable-rhythm model associates tone with syllable-rhythm.

With regards to ESL teaching/learning, Auer's research is worthy of note for two reasons: first, like Dauer's model, Auer's breaks down rhythm into features which could be used to compare the target language and the students' L1. Second, and more importantly, Auer's research goes further by examining the *raison d'être* of the different features and their relation. This represents a first step towards a hierarchical organization of phonetic and phonological traits that could be used in TESL. His results enable us to foresee a more sophisticated hierarchical system. For instance,

[14] Auer uses the term "shell complexity" to refer to "the maximal number of phonemes in the syllabic shell. Following Vennemann (1988), the shell of a syllable is defined as the union of the syllable's head and coda. The number of phonemes in the syllabic shell therefore equals that of the phonemes in the syllable as a whole, with the nucleus phonemes subtracted." Auer does not take into account the complexity of the nucleus in "shell complexity" because he considers it in the "segmental complexity".

while he found that certain processes are not directly linked to duration (see second point above), he found that others, namely, shell complexity, vowel reduction, and processes that enhance or destroy syllable structure are directly related to duration.

In fact, the results lead Auer to propose an alternative model that focuses on prosodic categories rather than on duration. This proposal takes its roots in the Nespor and Vogel's (1986) assumption *'that in any language there is a hierarchy of prosodic categories to which phonological and phonetic rules and regularities refer'* (Auer, 1993:90). The proposal is that languages have either the syllable or the phonological word as their basic category and are consequently labelled 'syllable languages' or 'word languages'. However, these names signify a tendency rather than a state; indeed, in the first case, a maximum number of phonological regularities and processes have as their domain the syllable, in the second case, a maximum number has as their domain the phonological word." (1993:91). As a result, Auer proposes two prototypical languages and their respective features, summarized in figure 3.2.

The prototypical syllable-language would have the following features:

- Dominant syllable related processes / phonotactics rules
- Syllable-enhancing processes:
  - Vowel epenthesis
  - External sandhi = internal sandhi
  - Allomorphs depending on syllable structure
- No vowel reduction
- No accent
- Tone Assigned to lexical syllables
- Low shell complexity C(G)...C or less

The prototypical word-language would have the following features:

- Dominant word related processes / phonotactics rules
- Syllable-destroying processes:
  - Deletion of vowels
  - Consonant-consonant assimilation
- Vowel reduction
- Accent
- Tone assigned lexically to phonological words
- High shell complexity CC...CC or more

Figure 3.2: *Features of the prototypical languages, syllable-language and word-language, proposed by Auer (1993)*

He then places the 34 languages examined according to these features. From this table (figure 3.3), one can observe that Mandarin is closer to the syllable-language prototype than to the word-language prototype; that French is halfway between the two prototypes; and that English corresponds to the toneless word-language prototype.

	Processes/ phonotactics	Syllable structure rules	Vowel reduction	Accent	Tone	Shell complexity
prototype	**S**	**S**			**S**	**L**
Yoruba	S				S	L
Amo	S	S			S	L
Navaho	S	S			S	L
Eskimo	S					L
Mundari	S	S				L
Vietnamese	S				S	L
Fijian		S				L
Basque	S/W	S/W		(+)		M
Japanese	S	S	(+)		(W)	L
Hausa	S	(W)	(+)		S	L
Toba-Batak	S	S		+		L
Mandarin	S		(+)	+	(S)	L
Quechua	S	S	(+)	+		L
Korean	S	S/W	(+)	+		M
Yidiê	S/W					M
Nimboran	S/W	S/W	+	+		M
French	S/W	S/W	(+)			H
Turkish	W	S/W		(+)		M
Telugu	W	W		(+)		L
Khalkha	W	S/W		+		L
Asmat	W					L
Uzbek	W	S/W	+	(+)		M
Tamang	W	W			W	L
Nama	W	S/W			W	M
!xóõ	W				W	M
Toda	W	S/W	(+)	+		H
Diegueño	W	S/W	+	+		M
Tzeltal	W	W		+		H
Klamath	W	S/W		+		H
Gaelic	W	S/W	+	+		H
Russian	W	S/W	+	+		H
Circassian	W	W	+	(+)		H
Arabic	W	W	(+)	+		H
English	W	W	+	+		H
prototype	**W**	**W**	**+**	**+ or W**		**H**

Figure 3.3 *Organization of languages according to their features (reproduced from Auer, 1993:93-94).*

Legend:
Processes/phonotactics rules:
W= Word related processes dominant
S= Syllable related processes dominant
S/W= both
Blank = does not apply or insufficient information

Syllable-destroying/ -enhancing processes:
W(ord)= Syllable-destroying
S(yllable)= Word-destroying
S/W= Both or none of them
Blank = does not apply or insufficient information

Reduced vowel system:
+ = overall reduction
(+) = marginal or restricted domain of reduction, or only [+/-long] neutralized
Blank = no reduction or does not apply

Accent:
+ = Word accent exists
(+) = Vague accent system
Blank = no accent or insufficient information

Tone:
S = Assigned to lexical syllables
W = Assigned lexically to phonological words
( ) = restricted tone language
Blank = mo or marginal tone

Shell complexity:
H(igh) = CC…CC or more (unless marginal or dialectal)
L(ow) = C(G)…C or less
M(id) = in between (incl. CG…CC)

More careful observation of the Mandarin features leads to the conclusion that this language is very much a syllable-language; indeed, the two criteria that distance Mandarin from the prototype are the existence of word accent and its classification as a restricted tone language. However these criteria were established because

Mandarin permits toneless syllables and, although this is true and legitimate, one must note that toneless syllables are relatively rare and that, except for these examples, the language is very close to the syllable-language prototype. French is a more complicated case as it has clear evidence of both tendencies. In fact, one can observe that French is syllable-based but with word-language tendencies. Indeed, various processes, such as syllable-destroying processes and shell complexity are associated with the word-language prototype. Auer's analysis that French is syllable-based with word-language tendencies – rather than word-based with syllable-language tendencies – seems to be explained by his comment that *Modern colloquial French [...] increasingly tolerates phonetically complex syllables [...]* (1983:10).

With regards to ESL, Auer's classification has three advantages; first, it enables us to see the phonetic processes used in a specific language. Second, if the phonological traits of a language permit a (synchronic) view of the place of a language on the scale, the changes in phonetic processes could reveal a (diachronic) perspective of that language. The third advantage is that it indicates much more clearly which processes two languages may share and which they may not share. This could obviously be useful for TESL as it helps pinpoint the areas that are most likely to be problematic for students; the latter would need to learn to activate processes that are not used in their L1 and de-activate processes that are used in their L1 but not in English. However, a same process can have different operating specifications or parameters in different languages. For instance, vowel reduction is a characteristic of Quebec French and of most varieties of English. In Quebec French the sound /i/ is laxed to [ɪ] when it occurs in a word final syllable closed by a consonant other than /v, z, ʒ, r/; in English however, [ɪ] is a phoneme or the reduced form of the vowels /eɪ, ɛ, i/ in unstressed syllables. Thus, while Quebec Francophone ESL students accurately produce words such as 'build' or 'pick' because it reproduces the conditions for the use of [ɪ] in their language, they often

mispronounce the underlined part of the following words 'happy', 'children', 'river', 'live'.

Thus, for a more complete and more useful system, it would be necessary to include the parameters of each process for each language. This obviously represents a colossal task especially in terms of establishing the phonological level at which processes take place, and, in case of potentially conflicting processes, the hierarchy to be observed. This is already dealt with, in part, by Optimality theory (discussed in the next chapter).

As we have seen, Dauer and Auer broadened the concept of rhythm by including features other than timing. However, the idea of using only duration, as Pike had originally proposed, was later revived, possibly because of new findings originating from studies that have investigated the measurement and the perceived location of speech events. One such revealing research was that of Morton et al. (1976) who found that the perceptual moment of occurrence of acoustic signals (in a syllable), what the authors called the "perceptual center" or "P-center", corresponds "neither to word onset, stressed vowel onset, nor to position of peak vowel intensity". This was seen as the possible explanation as to why it had thus far been impossible to find empirical evidence of the perception of rhythm and of isochrony. This led a few scholars to experiment with new methodologies.

3.2.1.4. Statistical approach to rhythm

Ramus (1997, 1999) proposes a new analysis of languages according to acoustic/phonetic measurements. The main objective is not to search for isochrony but to investigate the relative prominence between successive syllables. Rhythm is seen, not as a basic to the temporal organization of languages, that is to say, not as a property of language, but as a consequence of some of their phonological properties, and more specifically, of syllable complexity, correlation between syllable weight

and accent, and presence or absence of vowel reduction. Ramus' approach is based on the idea that the syllable complexity of a language can help to determine its rhythmic class and it relies on the idea that stress-timed languages, by allowing vowel reduction, should display greater vowel duration variability. He therefore measured the percentage of duration taken up by vocalic intervals (%V), the standard deviation of the duration of consonantal intervals within a sentence (ΔC), and the standard deviation of the duration of vowel intervals within a sentence (ΔV).

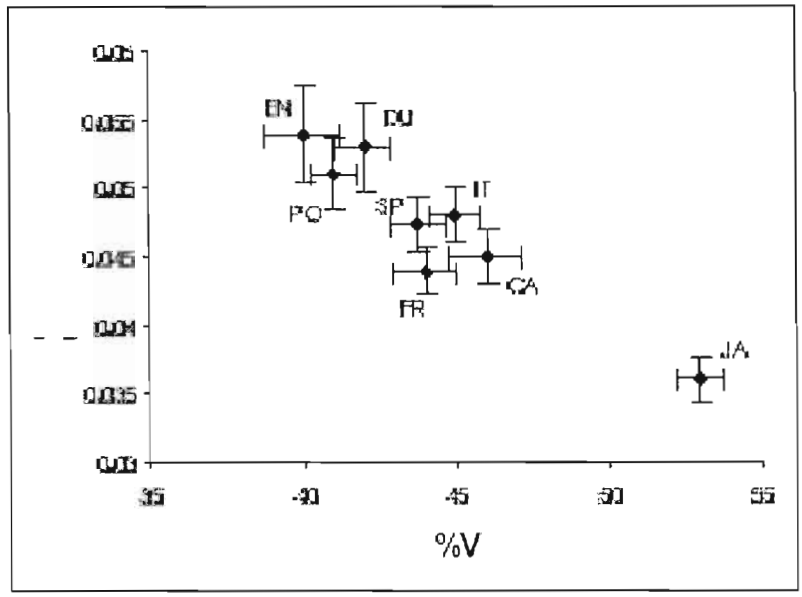


Figure 3.4 Standard deviation of consonantal intervals vs. proportion of vocalic intervals in 8 languages, 5 sentences times 4 speakers per language (reproduced from Ramus: 1999)

One of the most interesting findings of the research pertains to the standard deviation of consonantal intervals when compared with the proportion of vocalic intervals; indeed, on the graph (figure 3.4) English, Dutch, and Polish form one cluster, Spanish, Italian, French, and Catalan form a second and Japanese is clearly separated from the other two clusters. These clusters, which correspond respectively to the stress-timed, syllable-timed, and Mora-timed language groups, gave new hope to the stress-timing syllable-timing theory.

However, one can interpret the results differently. Ramus' hypothesis is that stress-timed languages have complex syllables and that the latter show a greater number of consonant clusters and thus that consonant/vowel ratio is greater. But when the author demonstrates that the greater ratios indeed – and unsurprisingly – correspond to languages including consonant clusters, he fails to show how this makes the languages stress-timed. In other words, and as stated by Cummins (2002), the experiment measures phonotactics of a language rather than rhythm.

> The discrete basis for the suggested taxonomy can be argued to be grounded in segmental inventories and syllabic phonotactics, and can therefore be accounted for without reference to anything resembling the pre-theoretical notion of rhythm described at the start of this section. More succinctly, where is the bom-di-bom-bom in %V? (Cummins 2002: 122)

In fact, as languages that permit consonant clusters also permit simple consonant sequences, it is easy to imagine that the experiment reproduced with different sentences would show a lesser difference between the groups. Furthermore, and as later recognized by the author (Ramus, 2002:115), it is possible that the measure reflects rhythmic differences but not rhythmic classes and it is highly possible that the analysis of more languages would fill the space between the clusters and show a continuum as suggested by Dauer's model.

Still, the clusters obtained cannot be so easily dismissed because they seem to offer evidence for Pike's dichotomy. However, the results might not fuel the arguments supporting the stress-timing syllable-timing hypothesis as much as those concerning human perception of rhythm. Indeed, it might just be that humans are able to segment and compute language into vowel and consonant ratios.

In addition, the analysis of the standard deviation of vowel intervals vs. proportion of vocalic intervals (figure 3.5) does not show any significant sign of classification. Indeed, one of the most striking points is that Polish (which clustered with English

and Dutch in the 'stress-timed' languages group) is detached from all other languages.

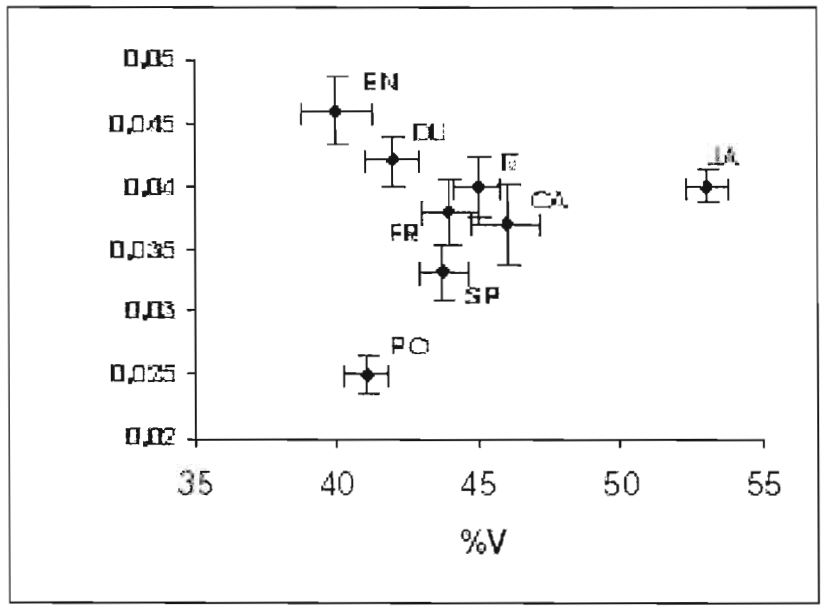


Figure 3.5 *Standard deviation of vowel intervals vs. proportion of vocalic intervals in 8 languages, 5 sentences times 4 speakers per language (reproduced from Ramus: 1999)*

This result, as explained by the author, highlights the fact that vowel reduction is not found in Polish and thus clarifies why some linguists consider Polish to be stress-timed and others syllable-timed.

In view of the opinions expressed here, it is logical to conclude that research based on acoustic correlates ratios are not expected to bring much to TESL. The point about perception of rhythm is important though in the sense that students have to learn to perceive English as much as produce it.

3.2.1.5. Pairwise Variability Index

Low and Grabe (1995) proposed a system for classifying differences in rhythm based on the Pairwise Variability Index (PVI), which, like Ramus' hypothesis, stems

from vowel duration variability. The PVI is calculated from the absolute value of differences in vowel duration between successive syllables, divided by the average duration of the pair; thus, one should expect higher PVI values in stress-timed languages because the durations of vowels in successive syllables vary much more than those of syllable-timed languages. The main difference between this metric and Ramus' is that it incorporates a normalisation component for speaking rate. Again, the results have not provided empirical evidence of the stress- syllable-timing hypothesis, but again, at least one of the authors is reluctant to drop Pike's theory altogether:

> The results have provided some support for phoneticians' classifications of languages as stress-timed or syllable-timed [...]. Unambiguous vindication of a categorical distinction between stress- and syllable-timed languages, however, has not emerged. (Grabe, 2002)

The results obtained in this research, which included 18 languages, indicate that the rhythmic differences among these languages cannot be expressed in terms of classes as no clustering is noticeable (figure 3.6)[15].

[15] Although the author does not specify it, it would seem that she attributed a clear circle to supposedly stress-timed languages, a dark circle to supposedly syllable-timed languages, a dark square to mora-timed languages and a clear square to languages of unidentified timing.

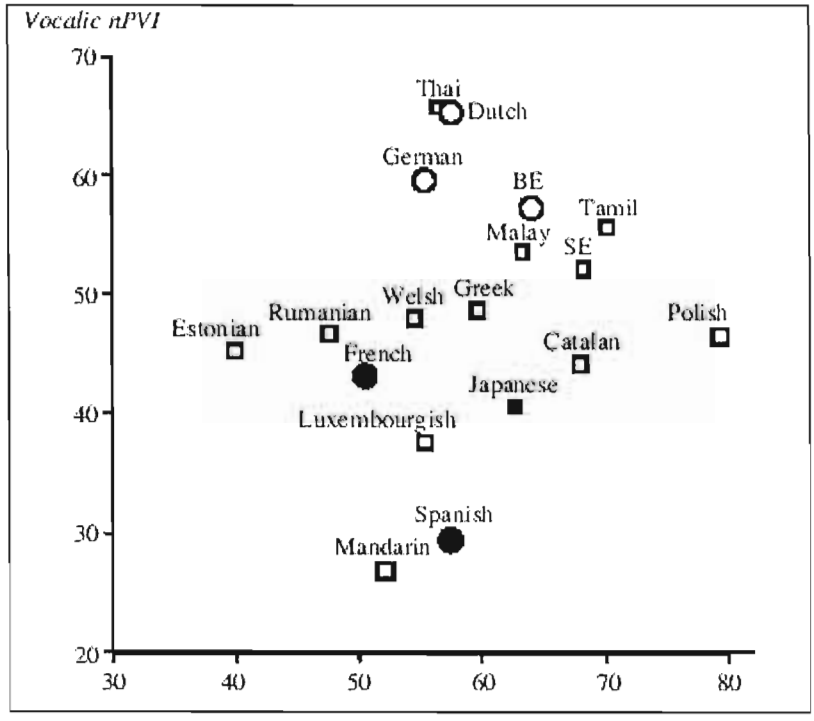


Figure 3.6 PVI rhythmic space for 18 languages. One speaker per language. 7 speakers from Spanish and French. BE= British English, SE=Singapore English. (reproduced from Grabe: 2002).

Furthermore, using the same PVI measures, Low (1998) demonstrated that dialects of a same language can obtain very different PVI values as British English was assessed as stress-timed but Singapore English as syllable-timed. In addition, Grabe (2002) specifies that establishing rhythmic typologies might be premature because the differences between speakers of a same language can be as great as those between different languages. This is illustrated in the findings of Pellegrino et al. (2002) of great rhythm variation between seven speakers of Spanish (figure 3.7).

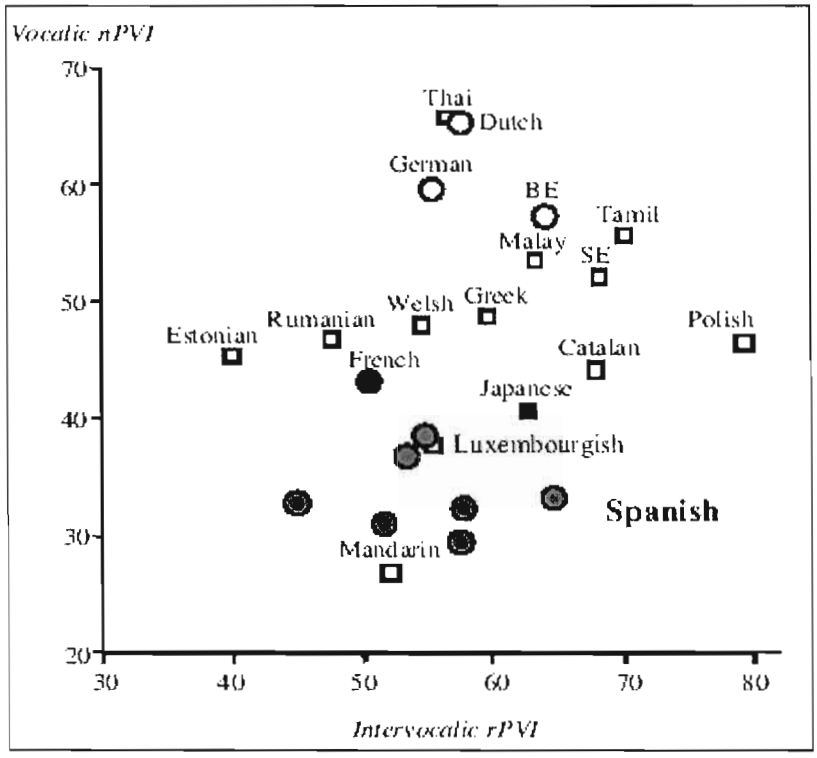


Figure 3.7 Same as figure 3.6 with the addition of seven speakers of Spanish (grey circles) (reproduced from Grabe: 2002).

These two examples of a statistical approach to rhythm are thus unlikely to provide the solution to the issue of classification of rhythm in the languages of the world and, some might say, they seal the fate of the stress-timing syllable-timing theory. For this reason, this approach should not be considered in the establishment of a classification system that could be useful for ESL. Its application, however, can be envisaged for analyses of cross-dialect and cross-speaker variation rhythm.

3.2.1.6. Speech cycling

A different and novel experimental method to study speech rhythm is *speech cycling* (Cummins 1997; Cummins and Port 1998; Port, Tajima and Cummins 1998; Tajima 1998; Cummins 2002; Tajima and Port 2003). The suggestion is that many phenomena such as 'stress', 'metrical hierarchy', 'foot', 'short vowel', 'stress-timing' should not be considered symbolic concepts but rather should be considered

manifestations of a general perceptual model. The experiment carried out by investigations of speech cycling involves the repetition of phrases such 'beg for a dime' and a full cycle is understood to be the duration between one occurrence of 'beg' and the next; the time of occurrence (*phase angle*) of the word 'dime' relative to the cycle is then measured. Cummins' (1997) results showed that when the subjects are asked to align the beginning of the phrase with the beat of a metronome, the word 'dime' occurred at 1/3 or 1/2 or 2/3 of the cycle; similar results were found for Arabic, as illustrated in Figure 3.8. [16]

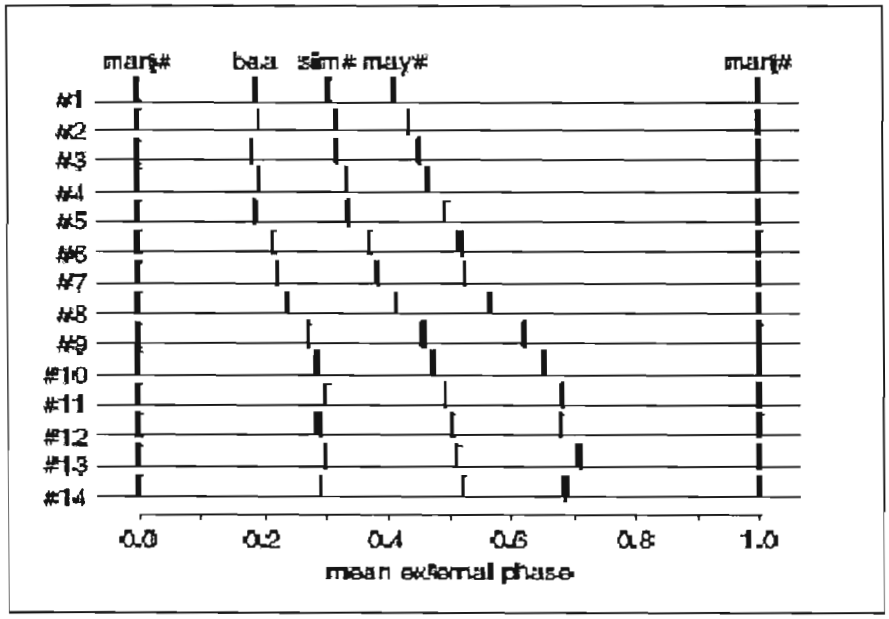


Figure 3.8 Sample measurements of syllable beats in external phase in Arabic. (Zawaydeh, Tajima and Kitahara 2002).

Tajima (1998) found that English speakers tend to place the stressed syllable halfway through the cycle while Japanese speakers tend to place the phrase final

[16] "External phase takes as the baseline the interval between the initial beats of the current and the following repetitions, and measures the time of occurrence of a syllable [...] relative to this baseline[...]. On the other hand, internal phase takes as the baseline the interval between the initial and final stressed syllable beats of a single repetition, and measures the timing of the medial stressed syllable beat relative to this baseline." (Zawaydeh et al., 2002:7)

syllable at the same phase angle. This leads the author to propose that the phase angle reflects what speakers consider to be linguistically important: in English, the last stressed syllable of the phrase, in Japanese, the phrase final syllable. Tajima also showed that, if a syllable is added to the left of the accented syllable, this accented syllable is displaced to the right and vice versa. This displacement is much greater in Japanese than in English, which could be seen as an attempt at isochrony in English.

The most important point of these investigations is that rhythm lies, not within a phrase, but at a higher level and has repercussions at the lower levels:

> By structuring an utterance so that prominences (stressed, beats) lie at privileged phases of a higher-level prosodic unit, rhythm is seen as an organizational principle which has its roots in the coordination of complex action and its effect in the realm of prosodic structure. (Cummins, 1998:167)

This finding is of interest in terms of TESL because it indicates for one thing that the emphasis should probably not be at the lexical or phrase level but at a higher level, for instance, the sentence. Furthermore, it would be worth investigating if ESL students would produce the repetition of a phrase such as 'beg for a dime' with the cycle of their L1 and, if so, if they could be led to learn the English 'beat'. However, speech cycling does not provide a classification system for languages. Zawaydeh et al. (2002) found that the speech cycling of Arabic is very similar to English, thus consolidating the generally accepted view that Arabic is a stress-timed language. This would seem to give credit to the stress-time syllable-time theory. On the other hand, the finding that English was more stress-timed than Arabic is in support of Dauer's view of a continuum.

Because of the repetition required for the tasks in these studies, the result cannot be assumed to be 'natural' sounding and could be refuted as evidence of isochrony; indeed, as illustrated in figure 3.8, the final phase angles are obtained around the

tenth repetition. However, Cummins and colleagues are not concerned by this unnaturalness as they discard the simple isochrony model of phonetics and seek empirical evidence of hierarchical organization claimed by phonologists.

3.2.2. Discussion and conclusion

The above examination of rhythmic classification has highlighted that, for the time being, no perfect classification system exists. The main conclusions drawn for the different propositions are summarized here:

- Pike's stress- syllable-timing theory, based on perception, does not stand to empirical scrutiny.
- Dauer's timing continuum's main flaws are that all features are considered equal and thus that two languages could share a place on the continuum without sharing a single feature. Furthermore, the author assumes that strong stress indicates stress-timing but the link between those elements has not been established, and in fact, is precisely what needs to be demonstrated. However, her proposition has the advantage of breaking up rhythm into features, which is interesting from the perspective of ESL research.
- Auer's proposal goes even further than Dauer's in terms of segmenting speech as the author takes into account phonetic processes, which, again, is valuable in terms of second language acquisition as it enables a comparison between L1 and L2 features. However, it was seen that the operating specifications of each process would also have to be considered (see the example of [ɪ] in Quebec French and English in 3.2.2.3). The sheer size of this undertaking, however, is probably enough to discourage its pursuit.
- The statistical approaches are less than promising, especially in terms of L2 acquisition research. Ramus' findings confirmed that a language with a high number of consonant clusters has a greater consonant/vowel ratio than a language with fewer consonant clusters. However, this measures phonotactics

and fails to prove the relation between consonant clusters and stress-timing, the author's premise. Low and Grabe's (1995) pairwise variability index gave similar results but also the potential for research in speaker and/or regional variation within a language.

- Finally, speech cycling offers a new perspective on rhythm: an organizational principle of which stress, consonant clusters, and other phenomena are manifestations.
- Speech cycling experiments on second language production could be interesting, first, to see if phase angle within a cycle differs greatly in the production by native speaker and by L2 speakers. Then, assuming the difference is considerable, to see if – and how – L2 speakers can be brought to adopt the native speakers' phase angle and the consequences this would have on their pronunciation.

The prevalent stress-timing syllable-timing theory remains the most evaluated theory but also the most discredited theory as it has found very little support at the empirical level, although the very recent speech cycling studies give it some credit. One of the reasons the theory might not be empirically confirmed is that it rests on the idea of isochrony, the measurement of which remains problematic:

- What are the events to be measured or measured from and which part of the events should be measured? We have seen that vowel onset and the perceptual center have been considered as well as the standard deviation of consonantal and vocalic intervals vs proportion of vocalic intervals.
- Also, while strict isochrony has probably never been predicted, how much deviation from it might be accepted? According to Allen (1975) speech timing is controlled with a variability factor of 4% (for long stretches of speech) to 10% (for longer segments). Lehiste (1977) states that decrements of 30 to 100ms were needed for listeners to hear differences of length in speech segments. So absolute isochrony is not needed for isochrony to be perceived.

- Finally what kind of speech should be used? Spontaneous speech, read speech or verse? Some scholars have used verses to show rhythm in speech; however, as highly structured rhythm is what differentiates verse from prose, it implies that prose is not naturally rhythmic, let alone isochronuous.

> The very existence of poetic metrics is justified by the supposed existence of a speech "without rhythm" as opposed to other kinds of speech "with rhythm", which is the raison d'être of normative metrics. (Bertrán 1999:107)

Generally, however and as we have seen here, isochrony is now assumed to be one of two things:

- An underlying constraint whose surface representation is imperfect because of phonological, phonetic, and grammatical features of language. "It is not necessary to 'prove' the phonetic existence of isochrony in order to accept it as a valid phonological concept." (Fox 2000:91)
- A concept related to speech perception, in other words, a "perceptual illusion".

In fact, facing the lack of empirical evidence of isochrony, some authors have concluded that rhythm itself is not production related but simply perception related:

> Consequently one is obliged to conclude that the basis for the distinction is auditory and subjective. A language is syllable-timed if it sounds syllable timed". (Roach, 1982:78)

It might even be seen as perceptual and accidental by some:

> "[...] the rhythmic differences we feel to exist between languages such as English and Spanish are more a result of phonological, phonetic, lexical, and syntactic facts about that language than any attempt on the part of the speaker to equalize interstress or intersyllable intervals" (Dauer, 1983:55).

It even led Cauldwell (2002) to replace Abercrombie's (1967) view of the stress-timing and syllable-timing hypothesis (which he refers to as "SSH") "*As far as is*

known, every language in the world is spoken with one kind of rhythm or with the other" with:

> We thus have to consider the possibility that production SSH can be replaced by a perception-SSH: 'As far as is known, every language in the world is perceived with one kind of rhythm or with the other ...' (Cauldwell, 2002:19)

A new perspective is now offered by speech synthesis, which provides evidence that irregularity of rhythm is necessary; indeed, synthetic speech, which observes strict rhythm regularity, is judged too regular by native speakers. Keller (2005), basing his research on Cummins and Port's (1998) experiments, showed that speakers (of French in this instance) introduce regularity and irregularity in specific places:

> [...] the degree of variability in the surface events depends on the strength of the vowel onsets. Close to strong vowel onsets (attractor sites), the variability would be small, while further removed from these sites, the variability could be much greater. (Keller, 2005).

Some authors even imagine that speech rhythm might be arrhythmic.

> If the aim of our investigations is to discover the relevant aspects of rhythm in language, one must include and consider all possibilities, from the existence of anisochronic rhythms (polyrhythmic structures), to the absolute lack of any kind of rhythm. (Bertrán, 1999: 126-7)

A total lack of rhythm is a rather disturbing notion especially as it would mean that we are completely tricked by our senses. We perceive rhythm. Infants and even primates have been shown to perceive the difference of rhythm between classes of languages (Ramus et al., 2000). So why is the existence of rhythm so difficult to prove? Maybe, as we have seen, the answer lies in the definition of rhythm: maybe rhythm does not need to be acoustically regular, at the phrase level at least. Also, from the results obtained from empirical research, one must concede that timing alone does not hold the key to rhythm and to the perception of it; in fact, very recent studies discuss "prosodic typology", which encompasses more than timing. For instance, Jun (2005), working in the framework of Auto-metrical phonology,

includes "word prosody", "postlexical prosody", and "timing" in her search of typology. Vaissière and Michaud (2006) sees prosody as accentuation (stress, tone, pitch accent, voice quality), intonation (syntactic and pragmatic), and as a performance factor; she comes to the conclusion that English is a stress language and French a non-tone, non-stress language. Rhythm is therefore absent from their concept of prosody.

We might also consider that, as the speech cycling experiments would indicate, rhythm is an organizing principle determined at a higher level of organization of speech. As phonology is the field concerned with the organization of speech elements, in the next chapter we will examine how different phonological theories deal with and explain timing and other prosodic properties.

Obviously, a major concern here is TESL and ESL learning and, in this respect, some elements discussed here have shown promise. For instance, Auer's (1993) introduction of rhythm features such as phonological and phonetic processes is valuable in terms of comparison of languages – and thus of students' needs; however, it was also noted, and illustrated with the case of vowel laxing in Quebec French and reduction in English, that his model is not sufficiently comprehensive as a same process can have different parameters in different languages. We will therefore search for a phonological theory to which these features could be anchored. Whether it will be at all possible to create a classification system out of this remains a question. As stated earlier, a classification system is desirable – which at least partially explains why the stress-timing syllable-timing typology has endured despite everything. However, we have to consider that finding a neat classification system might be an unrealistic ambition. After all, we don't expect to find categories of languages according to their segmental inventories. Trying to find rhythmic categories might the same as calling a language 'nasal' because it includes nasal vowels or 'fricative' because it makes use of more fricatives than any other

type of consonants. Although this might be true of the language, it does not suffice to describe it.

However, we ought to find ways to better understand and teach rhythm. As noted earlier, acquisition of rhythm is essential in ESL but, at the moment, little can be expected of learners when the information they are given is erroneous: "Speakers "time" their speech so that the amount of time between stressed syllables is about the same" (Lane, 2005:149) or "Syllables also make patterns that help us to figure out where words begin and end and which words are more important than others." (Dauer, 1993:83). The latter would make sense if rhythmic events, for instance peaks, were aligned with a specific place in the word, for instance, the first syllable.

> Compared to most languages, in which the stressed syllable is fixed, English is one for which stressed syllables are least likely to indicate word boundary. The following comment may therefore be very confusing to a student whose L1 stresses all content words – or possibly *all* words – equally. "English speakers use speech rhythm and intonation to show what's important. It would be confusing if all the content words were stressed equally." (Miller, 2000:65)

> From the following example, "Just as in music, English moves in regular, rhythmic beats from stress to stress – no matter how many unstressed syllables fall in between. This **stress-timed** nature of English" [...] (Celce-murcia et al, 1996:152)

It is clear that improvement of students' production of rhythm must start with the improvement of our understanding, closely followed by a much-needed revision of what teachers are taught.

3.3. Phonological organization of rhythm

From the previous section, we concluded that no prosodic classification model can be applied to second language acquisition in that sense that they do not permit to anticipate and or correct ESL speakers' prosodic errors. The phonological organization of rhythm for three languages English, French and Chinese will now be described. The premise is that phonological features of prosody might be transferred from L1 to L2. After a thorough description within the Autosegmental-Metrical framework, the three prosodic phonologies will be compared to highlight areas that would be favourable for prosodic transfer from French to English and from Chinese to English.

3.3.1. Phonological framework

In order to evaluate the existence, degree and nature of transfer in second language rhythm production, it is essential to define and describe the rhythm of each of the languages studied here. A phonetic description alone is not sufficient, as all phonetic features are not significant.

At this stage, the phonological organization of the three languages will be examined. The greatest challenge in this undertaking is that, phonology being the underlying and imperceptible sound organization of a language, different people, different times, and different schools of thought have brought their own version or vision. As mentioned by Fox (2000):

> However, it remains the case that there is no universal consensus among phonologists about either the nature of prosodic features themselves or the general framework for their description, and it is difficult to obtain a clear picture of the field as a whole. (Fox, 2000:1)

A phonological model should, in theory, be universal in the sense that it should be applicable to all languages. And all such models aim to be. However, some phonological theories seem to lend themselves better than others to the description

of certain languages. With regards to the present research, the framework used is that of Autosegmental-metrical (henceforth AM) analyses which have long been developed for English and applied, more recently, to French and, to some extent, to Chinese.

As mentioned by Di Cristo (2003), the AM framework has advantages:

> Cette approche bénéficie notamment de l'élaboration d'outils de formalisation qui sont applicables à l'ensemble des langues et qui permettent ainsi de définir, sur des bases comparatives plus rigoureuses, les traits saillants de chacune d'entre elles. (Di Cristo, 2003:4)

For each language, a summary of the history of the research on prosody will be given before the chosen AM analysis is described.

3.3.2. Autosegmental-Metrical Phonology (AM)

The term Autosegmental-Metrical phonology as applied to the study of intonation refers to the description of tones as autosegments and of their relationship as well as the way these tones are associated with the tone-bearing units in the text (the syllables) – the Metrical aspect. Indeed, in the AM framework, prosody is described in terms of high (H) and low (L) tones which correspond respectively to F0 peaks and F0 valleys. Some of these tones are associated with a syllable and others with prosodic unit boundaries. AM phonology is also based on the idea of a hierarchy of prosodic (intonational) units in which the Strict Layer Hypothesis (Selkirk 1986) entails that each constituent is composed of constituents of the level immediately below which components necessarily belong to the higher unit. The number and nature of these intonational components vary from one language to the next. This Strict Layer Hypothesis has been the subject of debate for some time because some constituents seem to include constituents two or three levels down. Shattuck-Hufnagel and Turk (1996:207-9) describe some of those cases.

The Autosegmental approach was first developed for tone languages. However, both tone languages and non-tone languages exhibit H(ighs) and L(ows) in their fundamental frequency. As specified by Duanmu:

> In generative phonology, the main difference between tone and intonation is functional, in that tone is used to contrast lexical meaning and intonation is used to contrast contextual meaning.[...] In this sense, both Chinese and English have tones, whether they distinguish word meanings or not. (Duanmu, 2004:6)

The Autosegmental-Metrical account of phonology was born out of necessity. Indeed, generative linguistics did not deal with tonal features but referred to Wang's 1967 theory, which was particularly complex. Finding that theory inadequate, Woo (1969) proposed that it would be more adequate to associate the feature "High" or "Low" to the segments – the adjacent vowels. In order to remedy this theory's own shortcomings – the exclusion of the possibility of tone on syllables containing a short vowel -, Williams (1976) and Leben (1971) designed their own melody mapping theories. In these, segmental and tonal information were separate at the beginning of the derivation but later reunified with a universal mapping procedure. The latter theories also had their flaws, in particular the specification with regard to the place in the derivation when the mapping would take place. (See McCarthy 1982 for more details). Although in these proposals the separation between segmental and tonal elements was only temporary, one could perceive the burgeoning idea of a multilinear phonology.

It is only with Goldsmith's (1976) proposal that the autosegmental aspect of phonology really took shape as it did away with the notion of the all-inclusive segment altogether by keeping the tonal and non-tonal features distinct throughout the analysis. This means that the different types of information are placed on different levels called "tiers", as can be seen below.

Segmental tier	pabatakala
Association lines	| | \/ /\
Tonal tier	H L H LH

The only way the elements of the two tiers are connected is with lines that show their association. As illustrated in this representation, a tonal feature can be associated with two segmental features and two tonal features can be associated with one segmental feature. It can also be noted that the lines do not cross. These stipulations about the connection between the tiers and about the absence of crossing formed what Goldsmith called the Well-formedness Condition (figure 3.9).

Well-formedness Condition (WFC)

a. Every tonal element is associated with at least one tone-bearing element; and every tone-bearing element is associated with at least one tonal element.

b. Association lines do not cross.

Figure 3.9 Goldsmith's (1976) Well-formedness Condition (from McCarthy 1982)

3.3.3. Phonology of North American English

Pierrehumbert (1980), in her influential doctoral dissertation, applied the model developed by Goldsmith (1976) to the description of English[17] intonation. This is the framework that will be used to describe English in this work and thus the model that will be described here. Relevant and substantial development by the same or other authors will also be presented to provide a more elaborate picture.

3.3.3.1. General phonological structure

In North American English, pitch patterns (i.e. strings of peaks and valleys) are associated with different prosodic units, levels of phrasing referred to as domains. In her dissertation, Pierrehumbert only recognizes one domain: the intonational

[17] While Pierrehumbert titled her 1980 thesis "The phonetics and phonology of English Intonation", it would appear that the variety she describes is American English. Indeed, her main speakers, named in the acknowledgments section (1980:3), are native speakers of American English, that is to say, the variety of English spoken in the United States.

phrase domain (IP) which corresponds to a sentence. She (Beckman and Pierrehumbert, 1986) later formulated the need for another accentual domain: the intermediate phrase (ip)[18]. These two domains have been labelled differently by different authors, as can be seen from figure 3.10.

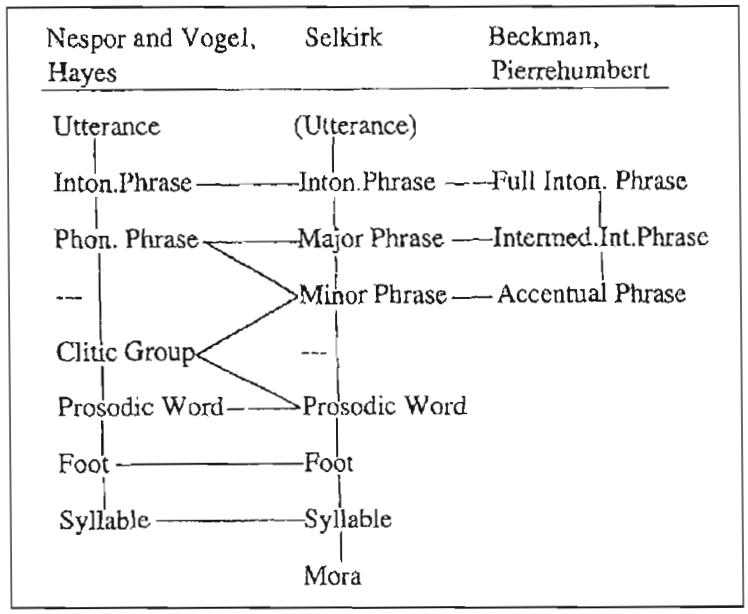


Figure 3.10: *Comparison table of prosodic constituent hierarchies and the terms used by different authors. From Shattuck-Hufnagel and Turk (1996:206)*

The additional level of phrasing (ip) in Beckman and Pierrehumbert's 1986 model departs from previous models of prosodic constitution, as noted by Xu and Xu (2005:3):

> Unlike the British tradition which generally recognizes only sense groups (O'Connor and Arnold, 1961), intonation-group (Cruttenden, 1997) or tone-unit (Crystal, 1969), the AM theory assumes that there are two levels of phrasing in English intonation: the intermediate phrase and the full intonational phrase." (Xu and Xu, 2005:3)

[18] Beckman and Pierrehumbert (1986) also identified the accentual phrase (AP) as a prosodic level of phrasing. However, while they found that the AP is essential in Japanese, they acknowledge that it is dispensable in English. "However, the evidence for the accentual phrase as a necessary unit in the prosodic hierarchy in English is much less definitive than for this phrase type in Japanese." (1986:270). For this reason, in this thesis, only IP and ip are recognized as domains of English prosody.

3.3.3.2. Accent types and contours

The tonal tier includes three types of tones: pitch accents, phrasal tones and boundary tones.

3.3.3.2.1. ***Pitch accents*** are tonal patterns assigned to strong syllables, that is to say, "stressed syllables". These accents can be monotonal (for example H) or bitonal (for example HL). The association of a tone with a stressed syllable is transcribed with a star. For instance, a single tone pitch accent is necessarily either H* or L*. Similarly, one of the two tones of a bitonal pitch accent will be marked with a star, e.g. H*L, HL*. Unstarred tones, marked with a raised hyphen H^- or L^-, in pitch bitonal accents can undergo spreading[19]. The alignment of a bitonal accent with the stressed syllable is marked in the same manner; for instance, H*L- indicates that the accent is high - low and that the high component is aligned with the stressed element of the segmental tier. Thus, the diacritics only mark the association of a tone with segmental elements and do not symbolize greater or lesser value. Accordingly, H* and H^- are both tones of a same H value and the former is not more H than the latter.

As there are two possible tones, that either tone can be single (necessarily starred) or paired up, and that either tone of a pair can be starred, one should expect 10 possibilities, as shown in figure 3.11 in which all possible tones with a H* are presented in the left column and all possible tones with an L* are presented in the right column. However, three of these (shaded in figure 3.11) are not actually realized. Pierrehumbert (1980:23) explains this by the fact that "the implementation of accents is such that there are no contexts in which they would be contrastive."

---

[19] Spreading refers to the fact that a unstarred tone (i.e. a tone not associated with a stressed syllable) can take on the feature of the previous starred tone. "The tone spreading rule [...] spreads T^- to the right when the next tone is phonetically equal or higher" (Pierrehumbert, 1980:76). For instance, "an L^- or H^- before H% spreads to the right, with the result that the F0 contour does not show a gradual rise from the phrase accent to the boundary tone, but instead a plateau followed by a sudden rise." (Pierrehumbert, 1980:219)

The distinction between the high accent (H*), and the two rising accents (L⁻H* and L*H⁻) was long disputed; for instance, Ladd (1983) contended that the difference between H* and L⁻H* was a delayed peak for the second one, which he calls the "scooped accent". However, Arvaniti and Garding (to appear) assert that their "results suggest that Pierrehumbert's three-way distinction between H*, L*H, and LH* is essentially correct, though it may not apply to all the dialects of American English."

H*	L*
H*+H⁻	L*+L⁻
H⁻+H*	L⁻+L*
L⁻+H*	L*+H⁻
H*+L⁻	H⁻+L*

Figure 3.11: Summary of possible pitch accents in American English proposed by Pierrehumbert (1980). Boxes of contours that are not actually realized by speakers are shaded. All possible tones with an H* are presented in the left column and all possible tones with an L* are presented in the right column.

3.3.3.2.2.***Phrasal accents*** are tonal patterns found at the end of a phrase, after the last pitch accent. These accents line up with the right edge of a phrase boundary and are made of one tone only: H or L. Because phrase accents, like unstarred tones in pitch accents, can undergo spreading, they are also marked with a raised hyphen: H⁻ or L⁻. Phrasal accents are relevant to mark the distinction between the accents identified in the British tradition as nuclear and pre-nuclear accents. Indeed, the AM theory does not distinguish between accents according to their place in the IP but according to their shape. Thus, what the British School considers the Head (Prenuclear accent units) and the Nucleus (Nuclear Accent) are, respectively, in AM, a pitch accent and a pitch accent followed by a phrasal accent.

3.3.3.2.3. ***Boundary tones*** are tones associated with the utterance as such. They are the final tones used to indicate if the final tonal pattern is H or L. To indicate its sentence final association, this tone is followed by the diacritic %.

Pierrehumbert (1980: 28) specifies that all four possible combinations of the two phrasal tone and the two boundary tones are possible and realized by speakers.

From all of these points, Pierrehumbert draws the grammar of intonation reproduced in figure 3.12. One can notice that the top line in the pitch accents inventory is a "return" line (i.e. it does not have a tone associated with it and goes backwards). This is to show that more than one pitch accent can be found in a phrase. In fact, the author specifies that two or three pitch accents are commonly found in phrases and that more than five is quite a rare occurrence.

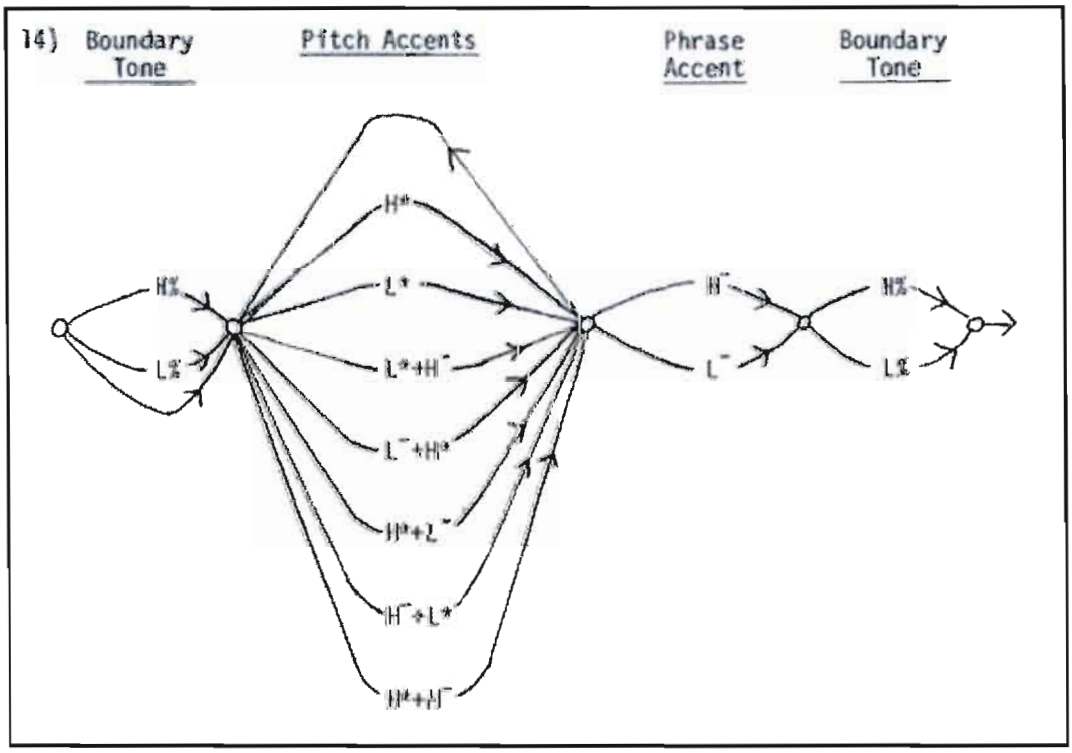


Figure 3.12: *Schemata of the possible tones and their association to form the grammar of North American English intonation. Pierrehumbert (1980:29)*

Pierrehumbert (1980:22) considers that:

> The well-formed tunes for an intonation phrase are comprised of one or more pitch accents followed by a phrase accent and then a boundary tone. (There is also a leading boundary tone after a pause). The pitch accents consist of either a tone or a pair of tones.

3.3.3.3. Association rules & tonal alignment

"Tonal alignment can be defined as the temporal synchronization of tones and segments, mediated by the prosodic structure." (D'imperio 2000:10). So far, we have seen that pitch accents are associated with stressed syllables. It must be noted however that "associated" does not necessarily mean "aligned with". Indeed,

> As Bruce (1977) made clear, the relationship between alignment and association is not always straightforward. That is, phonological associations cannot simply be read from the F0 curved. In many languages, for example, accent peaks often appear beyond the syllable or mora to which they are associated. Welby (2006:344)

The analysis and explanation of the discrepancy between phonological association and phonetic alignment are beyond the scope of this work. Xu and Liu's (2006) proposal of a "target approximation model" offers a new perspective on the issue.

For this work, tonal alignment will thus be limited to the "starredness" of a tone, denoting that the tone is phonetically aligned with a stressed syllable.
It will have to be borne in mind, though, that alignment is clearer at the phonological level than it is at the phonetic one.

The organization of stress at the word and phrase levels in English is of utmost importance and must be described. Indeed, stress is a property of lexical items and English rhythm is characterized by the distribution of stress in words and phrases. The rules that generate stress placement will thus now be summarized.

3.3.3.4.Stress Placement rules of English[20]: (bold indicates stress)

Monosyllabic words: The basic rule is that content words (nouns, verbs, adverbs) are stressed while function words (determiners, prepositions, etc.) are not. A more comprehensive account is presented below.

Stressed monosyllabic words:

- Nouns: ***cat, tree, car, mind***
- Verbs: ***eat, sleep, run, think***
- Adjectives: ***blue, slow, tall***
- Adverbs: ***out, fast, well, off***
- Numbers: ***one, two, three,***
- Demonstrative pronouns: ***this, that, these, those***
- Possessive pronouns: ***mine, yours, theirs***
- Interrogative pronouns (direct question): ***why, what, where.***
- Negation: ***no, not, none***
- Negative auxiliaries *I* ***can't*** *sing.*
- Auxiliaries used without a verb: ***Yes,*** *I* ***do.***

Unstressed monosyllabic words:

- Determiners: ***the, a, some, my, their***
- Personal pronouns: ***I, me, he, it, they***
- Prepositions: ***in, on, for, through, of***
- Demonstrative adjectives: ***this, that, these, those***
- Possessive adjectives: ***my, your, their***
- Conjunctions: ***and, or, but, which, where, that***
- Auxiliary when positive: ***can, do, have***
- Linking verb: ***to be***

[20] These rules and examples have been compiled over a number of years by the author of this thesis. For more details, see Selkirk (1982) and Dauer (1993) for instance.

- Most two-syllable nouns are stressed on the first syllable
    - "**mo**ther", "**jour**nal", "**ta**ble" and "**win**dow"
    - Some exceptions are found in basic vocabulary: "mis**take**", "sur**prise**", "de**sign**", "be**lief**".
- All other types of two-syllable words are stressed on their root.
    - di**stract, di**gest, ex**treme**, **le**nient, ar**ound**, **for**ward.

- As for longer words, the place of lexical stress is generally dictated by the suffix of the word.
    - Words ending in *–ese, –ee, -eer, –ette, -esque* and *–ique.* Are stressed on the last syllable.
        - Examples: Japan**ese,** employ**ee,** engin**eer**, cigar**ette**, pictur**esque**, volunt**eer**, Martin**ique**.
    - Words ending in *ize*, and *ate* are stressed two syllables before the suffix
        - Examples: a**po**logize, **ca**pitalize, **o**perate, pro**cras**tinate.
    - Words ending in *-tion, -sion, -cian ,-cial, -tual, -tial, -ious, -tious, -geous, -gious, -ical, -ity, -logy, –ify, -ic, itude,* which represent the majority of long words, are stressed on the syllable that precedes the suffix. This means that the stressed syllable is the penultimate syllable when the suffix is monosyllabic but on the antepenultimate syllable when the suffix contains two syllables:
        - Examples: fi**nan**cial, mus**i**cian, re**li**gious, **elec**trical, ac**ti**vity, as**tro**logy.

    - Certain suffixes do not affect stress placement. This is the case of *-able, -al, -en, -er, -est, -ful, -ing, -ish, -ism, - ist, -less, -ment, -ness, -ous, -y* and inflectional endings *(-ed, -(e)s)*

Example: **ma**nage, **ma**nage*able*, **ma**nage*d*, **ma**nage*r*, **ma**nag*ing*, **ma**nage*ment*.

A monosyllabic noun is de-stressed when it becomes the second element of a compound:

- **Black board** : both the modifier and the noun are stressed
- **Black**board : the modifier is stressed but the noun is unstressed.
- **Hot dog** : both the modifier and the noun are stressed
- **hot**dog : the modifier is stressed but the noun is unstressed.

- The stress may also shift from the ultimate or even penultimate syllable to an earlier one to avoid stress contiguity.
  - The girl is thir**teen**. The **thir**teen **girls** left early.
  - Abso**lu**tely. You're **ab**solutely **right**.

Pitch accents occur on stressed syllables, but not all stressed syllables necessarily carry a pitch accent. In fact, Pierrehumbert (1980:37), from the grid representation of metrical strength, devises the following rule of stress assignment (figure 3.13). The "foot" refers to the metrical foot which, "as discussed by Selkirk (1980) and Hayes (1980) is comprised of a stressed syllable and associated unstressed syllables" (Pierrehumbert, 1980:23).

22a) If a foot has a pitch accent, any foot of equal or stronger metrical strength in the phrase also has a pitch accent

except that

b) There are no pitch accents after the nuclear stress of the phrase.

Figure 3.13 *Pierrehumbert's rule of stress assignment (1980:37)*

She illustrates this with the following examples:

The nuclear stress of "California" is on the syllable "for" so a pitch accent is on "for" and another is possible on "ca". On the other hand, because the nuclear stress of "Abernathy" is on the first syllable, "A", the only pitch accent is that on this syllable and the syllable "na" cannot bare a pitch accent. Figure 3.14 illustrates the difference of stress assignment between the two words.

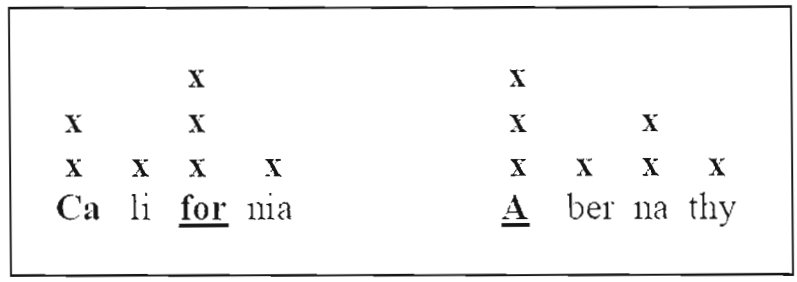


Figure 3.14: *Difference of stress assignment between two "California" and "Abernathy". Both the syllables "ca" and "for" have a pitch accent because the nuclear stress is on the later syllable. In "Abernathy", only "A" has a pitch accent because it is the nuclear stress and no syllable after it can have a pitch accent. Key: Light blue = unstressed; dark blue= stress/pitch accent; dark blue and underline= nuclear stress*

She also stipulates however that "if a word with this stress pattern occurs in prenuclear position, it can carry two pitch accents". The example provided is the word "perambulating" whose only pitch accent is the nuclear stress "am". Like in the case of "Abernathy", the potential other pitch accent cannot be realized because it occurs after the nuclear stress. However, when this word is used as a modifier of a

noun, as in "perambulating Peter", then both "am" and "la" can carry a pitch accent because the word is prenuclear to the nuclear "Peter". It therefore no longer has a nuclear stress but two pitch accents. Figure 3.15 illustrates the difference of stress assignment in this word in the two contexts.

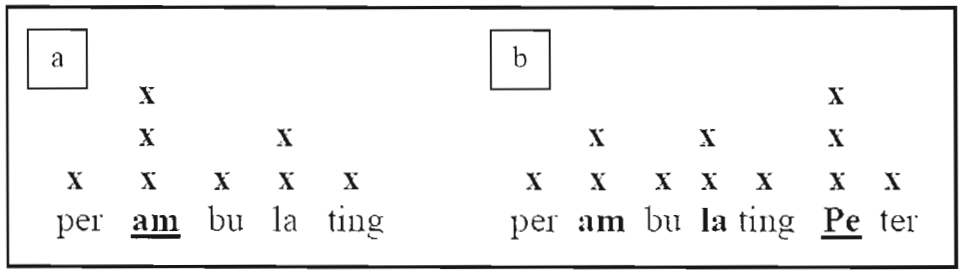


Figure 3.15: *Difference of stress assignment between "perambulating" used in isolation (a) and as modifier (b). In (a), the word has only one pitch accent because this pitch accent occurs on the nuclear stress. In (b), the word has two pitch accents because the word no longer has a nuclear stress – because it is in a prenuclear position.*

The author also concludes that a stressed syllable occurring after the nuclear accent does not carry a pitch accent but instead takes its contour from the phrase accent. She later takes the idea further by stipulating that this syllable not only emulates its contour from the phrase accent but may be the actual phrase accent carrier. She indeed specifies that:

> "[...] the phrase accent is not aligned with the metrically strongest syllable available. Instead, it is positioned with a fair amount of variation near the end of the word with the nuclear accent."(Pierrehumbert, 1980:44)

3.3.4. Phonology of French

For a long time, the rhythm of French was barely studied, as it was generally considered that there was no rhythm per se in French. The reasoning was that the only pitch movements in French were at the end of phrases and at the end of sentences and thus simply reflected intonation and syntax.

Only later was it understood that French had both rhythm and intonation markers but that they were fairly difficult to differentiate as they occurred on the same syllable:

> La caractéristique culminative particulière de la distribution de l'accent dynamique en français fait que tant l'accent tonal (T*) que le ton démarcatif de frontière intonative (T%) se réalisent phonétiquement sur la dernière syllabe des syntagmes intonatifs. (Cedergren et al., 1989:29)

The earlier belief that French had no rhythm necessitated that no other pitch movement be present where it could not be justified by syntax or intonation. To make this theory fit, many argued that the presence of a pitch movement at the beginning of a word was poor pronunciation, regional variation or emphatic stress, and, in any case, distasteful and improper. (See Vaissière 1991:101 for details)

The turning point towards the acceptance of a prominence earlier in the word or sense group in French was probably Fónagy and Fónagy's 1976 article which tactfully presented the phenomenon as early evidence of impending and probable change in French prosody.

With the acceptance of the possibility of rhythm in French, a number of theories and phonologies emerged. In the following section, the general guidelines of French phonology as well as a specific model will be presented. This section is titled "French of France"[21] because it describes research done on the variety of French spoken in France. This form of French has been the subject of the vast majority and the most exhaustive research on the prosody of French. Quebec French, on the other hand, has been comparatively little studied. The section on French of France will thus also be used as the basis for the next section, "Quebec French".

---

[21] The term "French of France" is here preferred for its political and historical neutrality compared to the term "metropolitan French". Other popular terms, such as "Continental French" and "European French", fail to recognize the differences, particularly with regard to prosody, between the varieties of French spoken in France, Belgium, Switzerland and Luxembourg among others. The variety of French referred to in this work is the standard variety of French spoken in France and used in many countries to teach French to non native speakers in many countries.

3.3.4.1. French of France

The following remarks are generally accepted in all current accounts of French phonology. These points aim at providing the reader with fairly basic information and description of the phonological situation of French as understood at the moment. The model adopted here, Jun and Fougeron (2002), will be described in detail in the next section together with a comparison, when relevant, with other current theories.

French differs from Latin and other Romance languages in that its stress domain is the phrase rather than the word. French is also characterized by rising prominence to demark these phrase boundaries. In this sense, stress does not have a distinctive function and is fixed on the last full syllable of the word (i.e. a syllable in which the vowel sound is not the schwa[22]) but is realized only if the word is the last of a phrase.

The main stress in French is marked by an F0 rise and takes place at the end of a prosodic unit. This prosodic unit has been called, among others, "rhythmic unit" by Di Cristo and Hirst (1993), "*syntagme prosodique*" by Vaissière (1991), and "intonation group" by Mertens (1993). In French, such a unit generally denotes a syllable bearing the main stress and potentially preceded by one or more unstressed syllables. This stress is called "main stress" or "primary stress" (referring to the rank of this stress), or "final stress" (referring to its position in the prosodic unit).

The other stress, also an F0 rise, is called "secondary stress" (referring to the rank of this stress) or "initial stress" (referring to its position in the prosodic unit). This latter term is somewhat misleading as the secondary stress rarely occurs on the initial syllable. However, when it does occur, this optional stress is produced in the first few syllables of the unit. Conversely, for Di Cristo and Hirst (1993), whose

[22] A distinction should be made between the English schwa and the French schwa: the former is a central non-rounded vowel used in reduced, and therefore unstressed, syllables. The French schwa is also central but is a little more rounded (although not as much as /oe/, the stressed form of the sound). In the definition given, schwa is understood to be the sound of the last syllable of "table", "Amérique", "alpiniste". However, for the vast majority of francophones, these final syllables are not pronounced at all.

conception of the prosodic unit (the Rhythmic Unit) includes sub-units (called Tonal Units), this other stress occurs at the end of a Tonal Unit. The tonal unit might well be limited to the first syllable of a longer word, the rest of which will be another Tonal Unit. In this sense and as can be seen from figure 3.16, for these authors, there are no secondary stresses (⬀) but only primary stresses (↗).

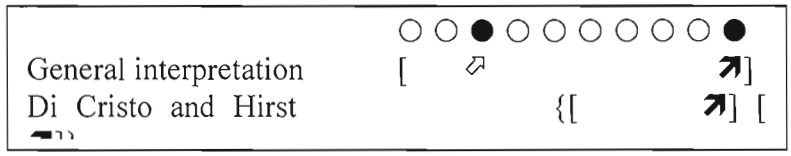


Figure 3.16 *Schematized representation of Hirst and Di Cristo's interpretation of French phonological segmentation.* {} = Rhythmic Unit; [ ] = Tonal Unit; ⬀ = secondary stress; ↗ = primary stress

Possibly because this secondary stress is not always realized on the same syllable, a number of authors have argued that what is produced is not the usual kind of secondary stress, whose function is generally understood to be the regulation of rhythm in strings of syllables (for instance, in long words in English) where too few primary stresses occur. Instead Rossi (1985) proposes that this stress is one of three different types of secondary stress. These types, explained in Lacheret-Dujour and Beaugendre (1999:44), include "melodic ictus" or "rhythmic accent", which comes into action to avoid stress clash. For instance, in *chaton gris*, the word *chaton* would bear a secondary a melodic ictus on the first syllable to avoid having two successive stresses, one on *-ton* and one on *gris*. The "enunciative stress" or "accent de focalisation" is the first stressable syllable of a prosodic unit unless its first sound is a vowel in which case it is the second syllable. The third type of secondary stress proposed is of the emphatic kind.

3.3.4.1.1. General phonological structure

The phonological model adopted for this research is the autosegmental-metrical account of French phonology presented by Jun and Fougeron (2000, revised 2002).

The authors propose two intonational units for French: The intonational phrase (IP) – the sentence – and its subcomponent, the accentual phrase (AP).

Description of the IP: The IP contains one or more APs. The final AP's final tone is marked with % to show the IP boundary tone.

Description of the AP: The AP contains at least one word of one syllable which bears the main stress. Each AP necessarily ends with a main stress, the final tone. This final tone is starred (eg. H*) to mark that it is AP final. The AP can also enclose a secondary stress, the initial tone, marked Hi. The difference between Hi and H* is thus one of position in the AP, not one of value or quality.

3.3.4.1.2. Accent types and contours

Jun and Fougeron's model (2002) describes the French Accentual Phrase to consist of LHiLH* in which LHi is the early rise and LH* the late rise (figure 3.17).

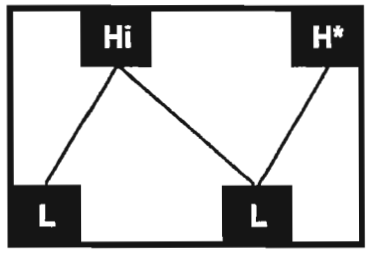


Figure 3.17: LHiLH*: The basic AP tonal pattern according to Jun and Fougeron's model.

The slope between Hi and the first tone of LH* depends on the number of intervening non-tone (i.e. unstressed) syllables (figure 3.18)

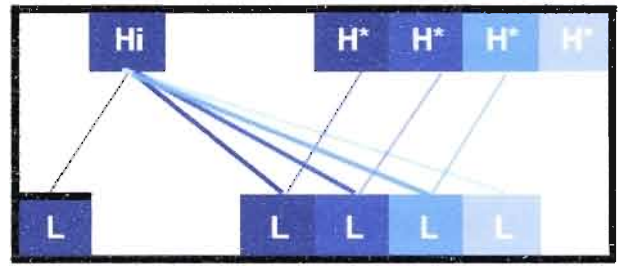
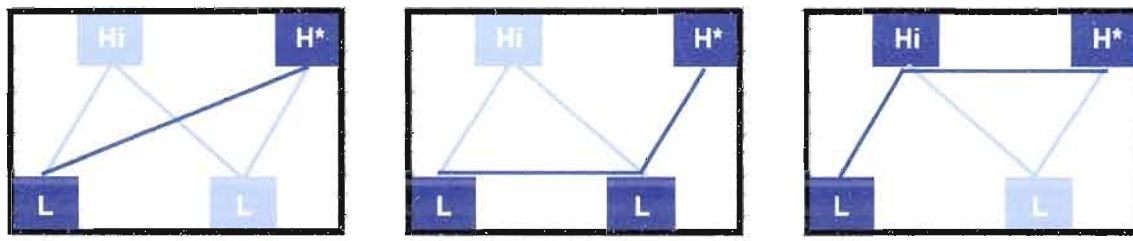


Figure 3.18 : Difference of slope according to the number of syllables between the early and the late rise. The greater the distance between the rises (represented here with the lighter colours), the gentler the slope.

Although LHiLH* is the most frequent accent realization, the AP tonal structure might be reduced by the non-realization of one or two tones, as explained in this next section.

Most often, the reduction of the AP tonal structure is achieved by the non-realization of one of the centre tones: Hi (resulting in the contour L~~Hi~~LH*) or the late L (resulting in LHi~~L~~H*) or both (resulting in L~~HiL~~H*). All of these productions are common for short APs (fewer than 4 syllables) but LH* is the most common one. Figures 3.19 a-c represent these three contours. For each figure, the light colour indicates the element(s) of the basic tonal pattern that are not produced. The dark line indicates the ensuing tonal contour.



(a) **(b)** **(c)**

Figure 3.19: *Common production of short AP in French. 19(a), the most common realization of the three, represents the ellipsis of both Hi and the first tone of the LH* accent. In 19(b) and 19(c) the Hi and final tone L are respectively not produced.*

On some rare occasions, the peripheral tones are not realised (see figures 3.20a-b). This rarity is interpreted by the authors as a tendency in French to preserve peripherals.

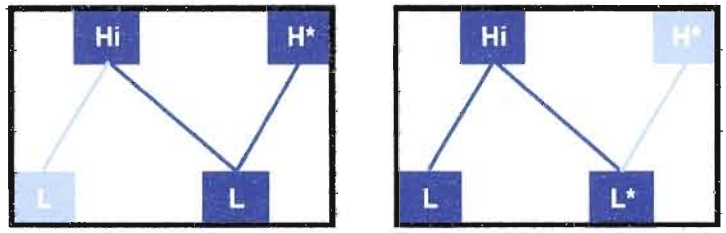


Figure 3.20: *AP tonal patterns without the early L (a) or the H* (b). As the late H is not realized, the previous L becomes starred.*

Despite this PRESERVE PERIPHERAL constraint, the initial L might not be realized if the first word is a content word and the AP is not the initial AP in the IP. The non-realisation of the H* is even rarer, as, according to Jun and Fougeron (2000), it would seem to occur only when H* would otherwise be flanked by other Hs, which, according to the authors activates the AVOID HHH constraint. In other words, in [...HiH*][Hi...] the H* would become an L and produce [...HiL][Hi...]. This should indeed be quite rare as it goes against the PRESERVE PERIPHERAL constraint. If this is a constraint, one has to assume that it is not of a very high rank as other solutions, which do not involve sacrificing a starred tone, would be possible. Indeed, in the example the authors give (see figure 3.21 and below), the noun group and the adjective are identified as two APs ([ ] mark AP edges and { } IP edges).

(2) {[Le garçon] [coléreux] [ment à sa mère]}
{[L Hi L*] [Hi L H*] [Hi L L%]}

However, it would be possible to reinterpret the recording as:

(3) {[Le garçon coléreux] [ment à sa mère]}
{[L Hi L Hi L H*] [Hi L L%]}

This latter interpretation of "le garçon coléreux" doesn't seem to fit with the basic tonal structure LHiLH* as this AP includes an additional LHi; LHiLHiLH*. However, the authors admit that, in some circumstances, tones may be added. Indeed, Jun and Fougeron (2002) specify that added LHi tones are normally found in cases of words of more than 7 syllables or long strings of clitics rather than in APs such as this one: "When an AP is longer than six syllables and contains two content words, the string will be produced in two APs with each content word forming on AP." (Jun and Fougeron, 2002:168). However, they also add that "Further data should be examined to confirm this observation." (Jun and Fougeron, 2002:168).

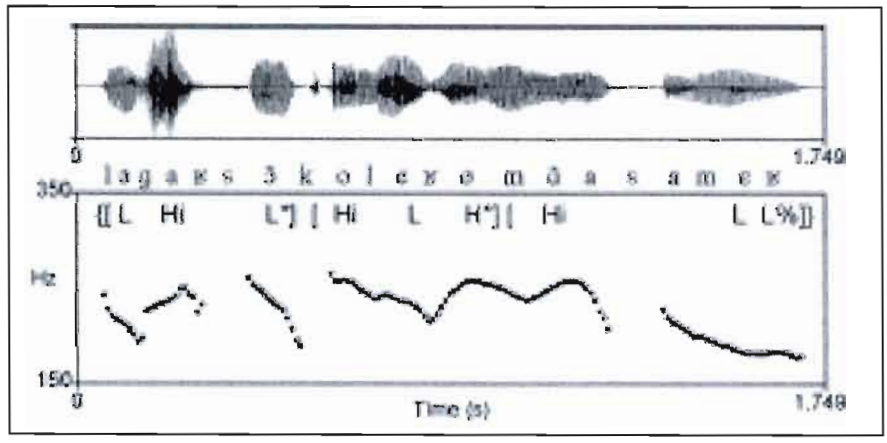


Figure 3.21: From Jun and Fougeron (2002:155, figure 4d) showing the authors' segmentation of "*le garçon coléreux ment à sa mère*".

3.3.4.1.3. AP domains and tonal alignment

The domain of the two tones (LHi and LH*) is the AP rather than the syllable. Indeed, while the initial L is linked to the beginning of the AP, it can spread over a number of non-tone syllables and the Hi might occur on the first, second, or, in some rare cases, the third syllable of the AP. In fact, Welby (2004:3) found that the early L is consistently realized at the function word/content word boundary and that the early H (i.e. Hi) is realized 46% of the time on the first syllable of the content word against 54% on the second syllable of the same word.

In fact, the results of her experiment

> "[...] strongly support the hypothesis that the L of the early rise is an edge tone (part of a compound edge tone) with a double association to the left edge of the first content word and to the left edge of the prosodic phrase or another syllable boundary." Welby (2004:3)

This seems to justify Post's (2000) use of "%L" to indicate the association of the initial L to the AP boundary and the early H.

Welby's research also indicates that there is no clear link between the early L and Hi and between the final L and H* as the timing varied between 50ms to 300ms for both tones. However her results did indicate that H* is predominantly (93%) produced late in the last full syllable and that 18% of the late L was produced on the same syllable as the H* and 82% on the previous syllable.

It all comes down to show that only the peripheral tones (the L of the early rise and H*) are anchored while the other two tones are not. This leads Welby (2006:343) to state:

> The French data thus provide evidence that the strong segmental anchoring hypothesis, in which both ends of rises have anchor points, cannot be generalized to all spoken languages.

3.3.4.1.4. Nature of the tones

The nature of the tones has been the source of disagreement between authors. For instance, Post (2000) claims that the initial tone is a pitch accent. With the finding that Hi is not associated with a particular syllable, this claim must be dismissed. On the contrary, the late accent (LH*) is associated with the final full syllable (i.e. not a "schwa" syllable). As maintained by Jun and Fougeron (2002), the final accent must thus be considered a pitch accent.

It is generally understood that the two tones of a bitonal accent must be linked. The two tones of the initial and of the final LH accents should be clearly linked. As we have seen however the tones can be adjacent or separated by a few syllables. Still, as specified by Welby (2006), the late L and its H* are never very far from each other. According to the same author, the role of the late L might be to contrast with, and therefore highlight, H*.

From the information gathered here, we can produce the following schemata (figure 3.22), similar to Pierrehumbert's (1980:29) for North American English, which shows the possible tones and their association to form the grammar of intonation of French.

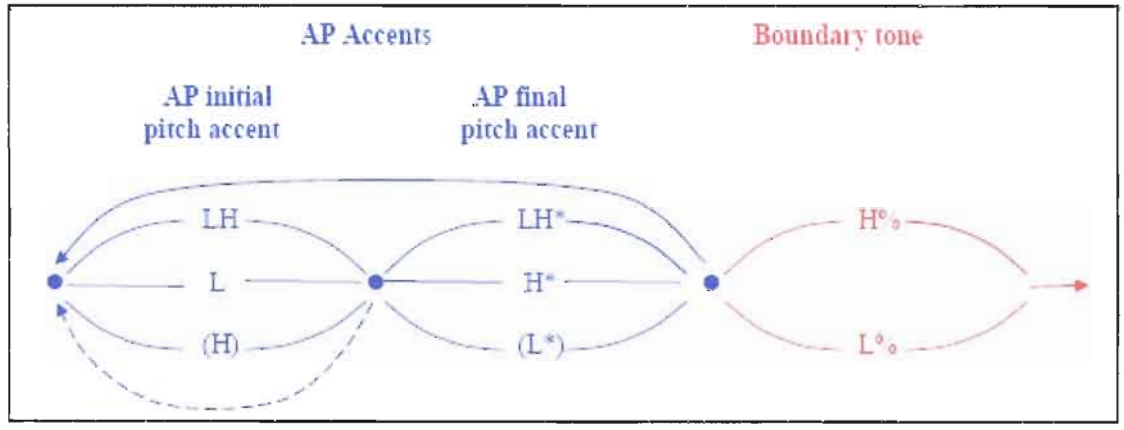


Figure 3.22: *Schema of the possible tones and their association to form the grammar of French. The brackets indicate that those tones are particularly rare. The AP initial return arrow (dashed) indicates that it may be possible to add a tone for particularly long APs.*

3.3.4.2. Quebec French[23]

In view of the location of this research, the majority of subjects are native speakers of Quebec French. The accent differences between French of France and Quebec

[23] The term "Quebec French" denotes the variety of French used in Quebec in everyday conversations in the media, schools, universities and businesses. This variety differs from but is mutually intelligible with French of France, which is considered "affected speech" in Quebec. It also differs from sociolects of the Quebecois working class such as "joual" which are more heavily accented and can be barely mutually intelligible with FF.

French (henceforth respectively FF and QF) is clearly noticeable even by non-native speakers of French at the lexical, syntactical, and phonetic/phonemic levels. QF prosody has been the focus of relatively few studies but has been subject of a few comparative studies with FF. These studies will be discussed here in order to anticipate whether the analyses for FF described above should be expected to be relevant for QF.

Ménard et al. (1999) showed that the markers Quebec francophone listeners used to perceive the linguistic origin (FF or QF) of a speaker were:

- Pitch height : FF speakers use higher pitch;
- F0 phrase contours: falling curves are associated with Quebec French while rising and flat contours are associated with French of France;
- Pitch range: Greater pitch range is used by Quebec Francophones than by France francophones. This perception-based finding was already attested by Bissonnette (1997) and confirmed in Kaminskaïa (2005:6).

Santerre (1990) found that in QF the stress clash constraint does not take place as long as the two Hs are of different nature.

> Voilà donc deux accents contigus, le premier de nature durative, le second de nature mélodique; il n'est pas nécessaire d'appliquer la règle de désaccentuation par contiguïté, parce qu'il n'y a pas de répétition consécutive du même procédé accentuel (Santerre 1990:49)

And thus that :

> Il y a donc bien une condition de non-contiguïté accentuelle en français de France, propre au français et au pékinois mais pas à l'anglais, selon Dell. Le moins qu'on puisse dire, c'est que cette condition n'est que facultative au Québec. (1990:55)

The alignment of the two tones of the final accent is reported to be the final stressed syllable and the preceding one:

> Le creux de la vallée mélodique se réalise généralement dans la syllabe précédente, mais dans les cas où la durée de la syllabe accentuée est

> longue, tout le complexe se réalise en synchronie avec la dernière syllabe du groupe intonatif. (Cedergren et al., 1989: 36)

This observation not only corroborates what was later reported by Welby (2006) as explained above but also provides a reason for the two tones of LH* to take place on the same syllable – a sufficiently long syllable – which in turn implies that the default setting would be for the two tones to be on the last two syllables.

Cedergren et al.'s data also confirmed the lengthening of IP final H* in comparison to IP internal H*. They attribute this asymmetry to the end-of-sentence lengthening phenomenon.

Kaminskaia (2005:2), who compares the phonetic realization of FF and QF prosody, found a difference in the production of the LLH* tone (bBH in her notation) by the two groups of speakers. Indeed, while FF speakers always produce a tone whose contour clearly resembles the diagram presented earlier (figure 3.19b, copied below), QF speakers produce a contour increasingly more modulated as the tone nears the end of the IP. Indeed, as can be seen from figure 3.23, in Kaminskaia's research, the LL part of the tone is flat for both FF and QF speakers (a) but is a falling tone for Quebec speakers (dotted lines) for the penultimate AP (b) and the ultimate AP (c).

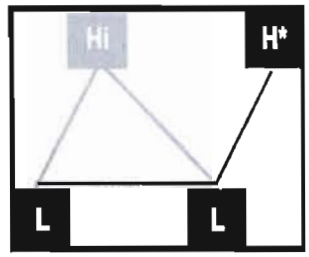


Figure 3.19b first presented on page 69

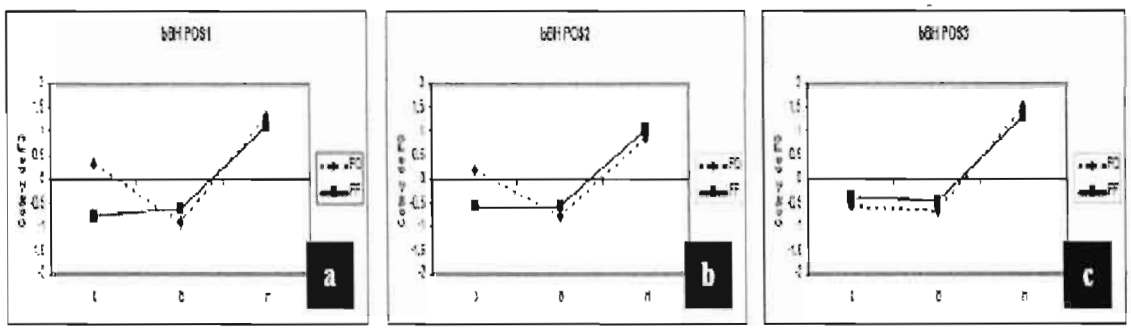


Figure 3.23: Reproduction of Kaminskaïa's (2005) figures 5, 4, and 3 respectively 23a, 23b, 23c and showing how FF speakers (plain line) produce the LL part in the same manner no matter the position of the AP while QF speakers produce a falling tone in the penultimate AP and even more so in the ultimate AP.

Furthermore, upon observation of these contours and their transitions, it appears, as explained by Kaminskaia, that after a H* tone, QF speakers aim for the final L rather than the initial L, while the FF speakers produce the first L, thus producing a flat passage between the two Ls. (Figure 3.24)

With a binary notation such as the one used in autosegmental metrical phonology, no difference is visible as, from the preceding H* to the second L, there is indeed a slope. From the extraction (figure 3.25) of the H*LL contours from Kaminskaïa's diagram, one can clearly see the difference between the concave pattern produced in FF and the convex pattern produced in QF. This observation leads to numerous questions such as: Is this difference perceptible to native speakers? If so, is it an important clue to identifying the speaker's variety of French? If that is the case, shouldn't more detailed notation be considered, at least for the description of finer detail such as that required in the description of varieties of a same language?

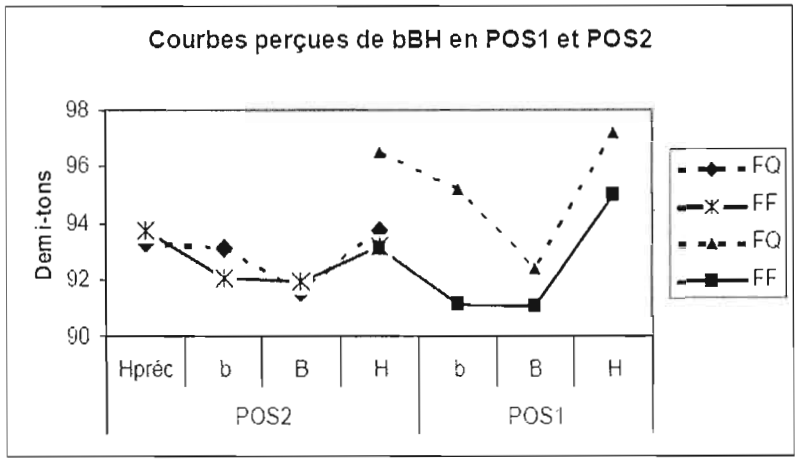


Figure 3.24: Comparison of QF and FF contours in the production of an LLH* AP (Kaminskaia, 2005: figure 11)

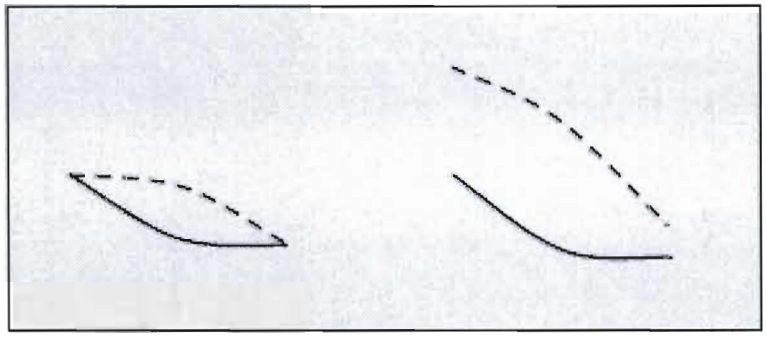


Figure 3.25 Contours of an H*BB string. QF contours (dotted line) are concave while FF contours (plain line) are convex.

Obviously, such considerations question the idea that targets, rather than movements, are important to detect, the founding principle of AM phonology. "[...] reaching a certain pitch level at a particular point in time is the important thing, not the movement (rise or fall) itself" (Bruce 1977:132)

3.3.5. Phonology of Standard Chinese[24]

As we saw earlier, AM phonology was originally developed for tone languages and later adopted for non tone languages. In recent years, however, a few authors (for example, Yip 2002) have chosen the Optimality Theory (OT) framework to analyze tone languages intonation. OT strength, as specified by Fox (2000:333), lies in the description and explanation of the mechanics of intonation. Indeed, this model posits a set of universal constraints which each language ranks differently, thus generating different outputs. In order to preserve a better prospect of comparison between the languages studied here, AM will be used for Standard Chinese (henceforth SC) as it was for North American English and for French. The difference between this investigation and the previous two however is that the research used is not an actual model. Indeed, while Standard Chinese phonology is getting more attention, no one has developed a model similar to what was proposed by Pierrehumbert (1980) and Jun and Fougeron (2000, 2002) for North American English and French respectively. Work by Peng et al. (2005) must be noted here as it is one of the rare works using the ToBI[25] framework for SC. It cannot be said to be faithful to the AM theory though, as it presents lexical and intonational tones on different tiers and adds a tier for stress levels of syllables (S3 to S0). The linearity of AM is thus compromised.

---

[24] The term "Standard Chinese" is preferred here because it is less historically and politically charged than "Mandarin" and "Mandarin Chinese" are; these latter terms also are ambiguous as they may refer to a category of Chinese dialects spoken in northern and south-western China or to one of these dialects – the one spoken in Beijing. The terms "Putonghua" ("the common language") and "Guoyu" ("national language") brim with exoticism but are somewhat discriminatory and intimidating to the non-sinophones. "Beijing Chinese" is too geographically limited even though Standard Chinese is based on the pronunciation of the Beijing dialect. The term "administrative language of China" would also be relevant as this form of the language was first used for education and administrative purposes. It would not however reflect that it has now been adopted by many as an every-day communication tool, that is to say, as "standard Chinese".

[25] The TOnes and Break Indices (ToBI) is a prosodic labelling system developed by Silverman et al. (1992).

For this work, the research of a number of authors will thus be compiled to offer a reasonably comprehensive view of prosodic phonology of Standard Chinese.

3.3.5.1.General phonological structure and prosodic hierarchy of Standard Chinese

Standard Chinese makes use of different pitch contours to allocate different lexical meanings to an item with the same segmental elements. SC is thus a tone language. It is sometimes assumed by non-tone language speakers that tone languages have neither stress, nor rhythm, nor intonation. Although a number of authors have now challenged this supposition, agreement on the features of these prosodic markers is lacking. The different views will be discussed after a description of SC lexical tones.

3.3.5.1.1.Tones

Table 1 displays the four SC lexical tones and the different systems used to catalogue them. Table 1 provides comparison of the graphic representation of the four tones.

Table 3.1 Comparative presentation of the notation of SC tones.

	AM notation	Tone shape	Chao tone digits[26]	Pinyin notation (on "ma")
1st tone	H	High	55	mā
2nd tone	LH	High rising	35	má
3rd tone	L(H)	Low falling (rising)	21(4)	mǎ
4th tone	HL	High falling	51	mà

[26] Chao tone digits are used to refer to the tonal space by means of numbers. The numbers are those found in the graphic representation of tones (see figure 3.26) in which 1 represents the lowest level and 5 the highest level. Tone digits 55 thus indicate that the tone (tone 1) starts high and ends high while 51 indicate that the tone (tone 4) starts high but ends low, making it a high falling tone.

Tone 3 is typically described and sketched as a low falling rising tone (as it is here in Figure 3.26). However, a number of studies has shown that tone 3 is actually a low falling tone (21) which can get a rising part in particular circumstances, such as in the citation form or in final position of an interrogative sentence (see Shen 1990, Lee 2005). This is why this tone is here described as an L (see Table 1) potentially ending in a rise (in brackets). Figure 3.27 thus offers a more accurate representation of SC tones as they are normally realized by native speakers.

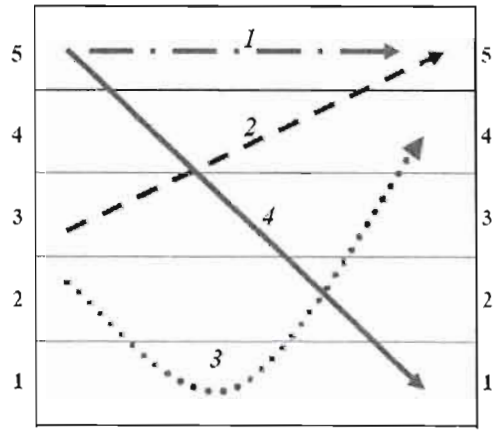


Figure 3.26: *graphic representation of the four SC tones with tone 3 as a low falling rising tone*

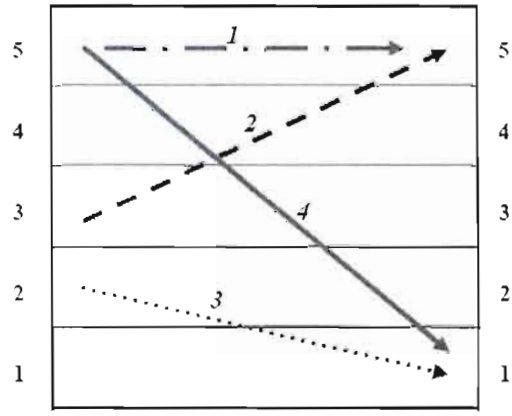


Figure 3.27: *graphic representation of the four SC tones with tone 3 as a low falling tone*

The adequacy of Figure 3.27 is confirmed by Kochanski et al's (2005:39) sketched outline of SC tones from their actual production. These modeled shapes of isolated tones are presented in Figure 3.28. For the purpose of comparison, the four frames have here been extracted and assembled into a single frame (Figure 3.29). The parts of the tones in the circled area clearly resemble the drawings in 3.27, including that of tone 3 as a low falling (without a rise).

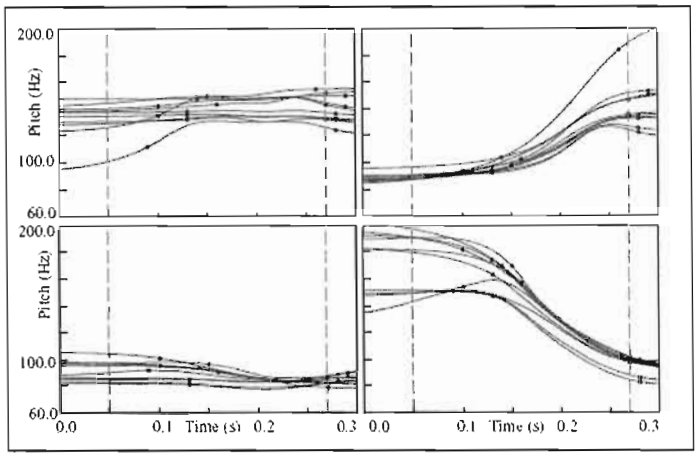


Figure 3.28: *Kochanski et al 2005 sketched outlines of SC tones from their actual production. The top left frame represents tone 1, top right tone 2, bottom left tone 3 and bottom right tone 4.*

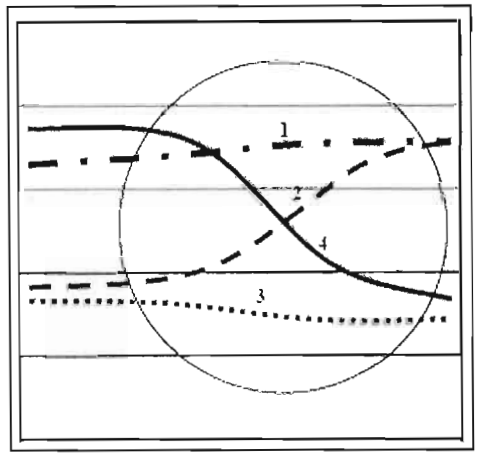


Figure 3.29: *Kochanski et al's sketched outlines of SC tones extracted and assembled in a single frame.*

SC also includes a neutral tone, sometimes called "fifth tone" or "zeroth tone". This tone does not have a defined pitch contour. It is phonetically marked by the light and short pronunciation of a syllable but, in terms of pitch height, varies according to the preceding syllable. Roughly speaking, the neutral tone is an L after the 1st (H), 2nd (LH), and 4th tone (HL) and an H after the 3rd tone (L). This tone – or this absence of defined tone - is for instance associated with the second syllable of some two-syllable words and with particles such as the structural particle "de" 的 used to indicated possession, the question particle "ma" 么 added to sentences to create an interrogative and the verb aspect particle "le" 了 which indicates the perfective aspect.

3.3.5.1.2.Tone rules

The tone associated with a lexical item may vary according to three rules:

• *Tone rule 1*: 3rd tone sandhi

When two 3rd tones are adjacent, the first becomes a 2nd tone as illustrated in figures 3.30a and 3.30b. The first figure sketches tonal sandhi between two third tones, drawn as LH sequences, that is to say, drawn as they are usually described but rarely produced. Figure 3.31 also sketches third tone sandhi but with the more realistic contour of the third tone as a low falling tone (L). For both figures, the left frame (a) shows the two third tones before tone sandhi and the right frame (b) shows that, with tone sandhi, the first third tone has been replaced by a second tone. Figure 3.32[27] presents the waveforms and F0 contours associated with 迋 (wǎng) "to go" which is pronounced with a third tone (low rising LH) in citation form but with second tone (high rising LH) when preceding another third tone. The low drop of F0 in the third tone valley reflects creaky voice, which is often used with the third tone in SC. (Keating and Esposito, 2007)

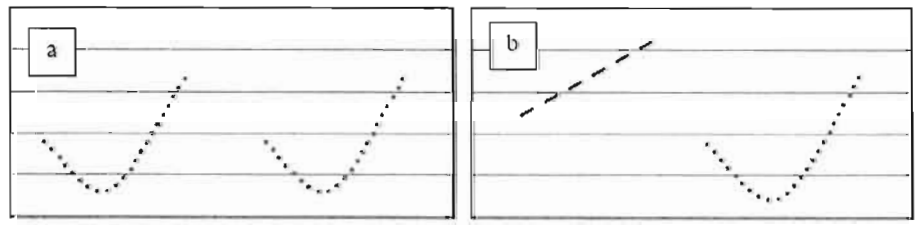


Figure 3.30: *3rd tone sandhi with third tone represented as an LH tone.*

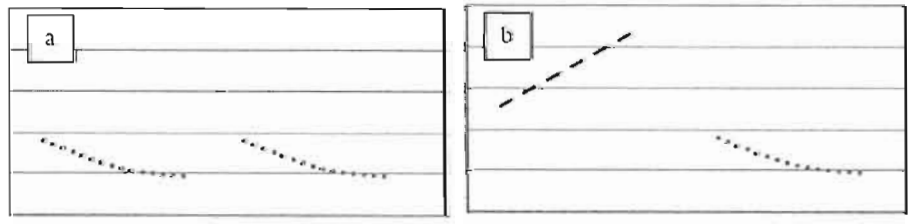


Figure 3.31: *3rd tone sandhi with third tone represented as an L tone.*

[27] All the spectrograms of Chinese presented in this thesis were done with data recorded by the Chinese subjects and were created with Praat.

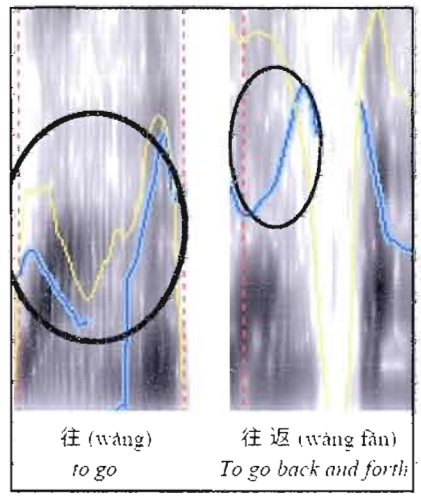


Figure 3.32: *Example of third tone sandhi with the word 迮 (wǎng) in citation form (left) and preceding another third tone (right) where it becomes a second tone.*

• *Tone rule 2*: Tone rule of the negation particle 不 (bù): the negation particle is produced with a 4th tone unless it precedes another fourth tone, in which case it is pronounced with a second tone (bú). An example is provided in figures 3.33a and 3.33b.

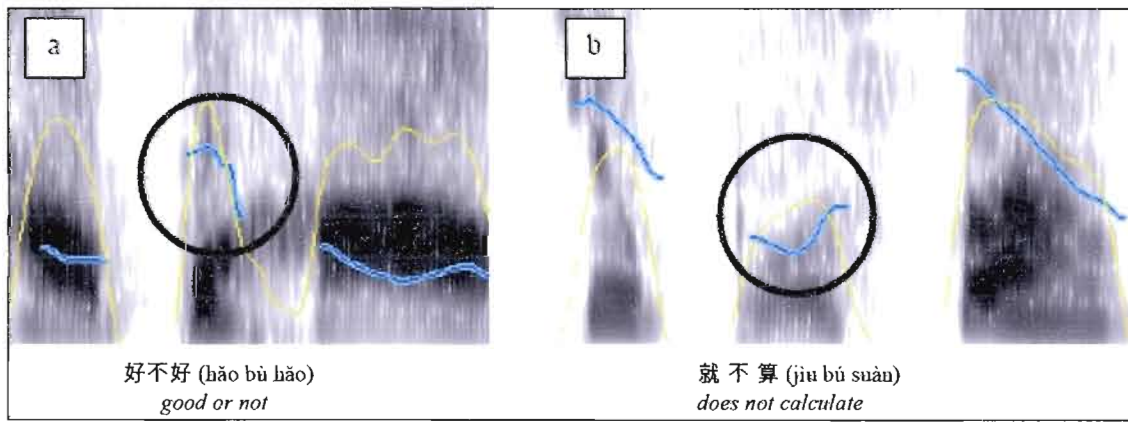


Figure 3.33: In circles, in (a) 不 produced with a fourth (HL) tone before a fourth tone and in (b) with a second tone before a second (LH) tone. (a): 好不好 (hǎo bù hǎo); (b) 就 不 算 (jiu bú suàn)

• *Tone rule 3*: Tone rule of the lexical sinogram 一 (yī): the word "one" is produced with the 1st tone when isolated or when used as part of a number. It is produced

with a second tone when it precedes a 4th tone and with a fourth tone when it precedes a 1st, 2nd or 3rd tone. An example is provided in figures 3.34a, 3.34b and 3.34c.

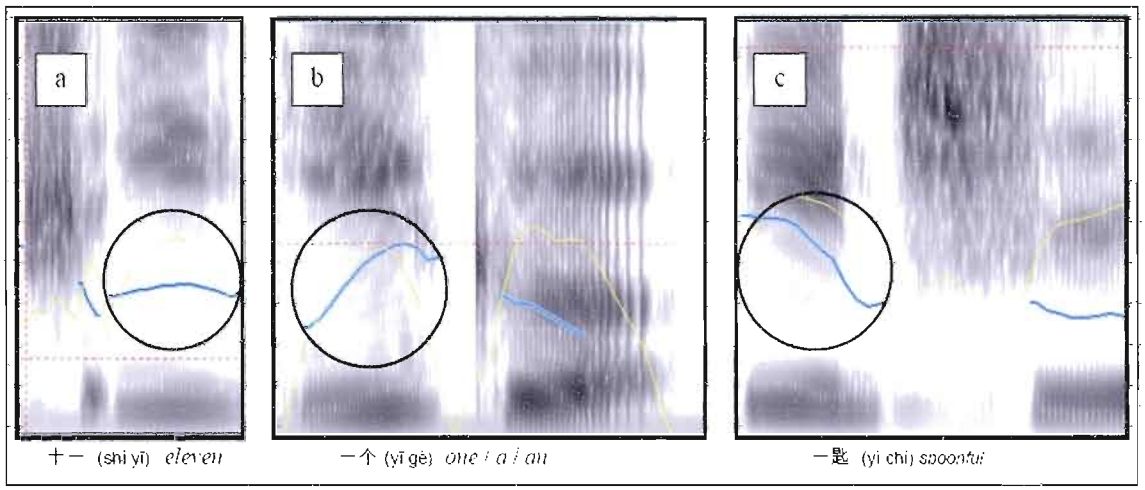


Figure 3.34: in circles, sinogram一 (yī), denoting "one", in (a) with a 1st tone because at the end of a number, in (b) with a 2nd tone because preceding a 4th tone and in (c) with a 4th tone because preceding a second tone. The examples are (a) 十一 (shí yī) eleven, (b) 一个 (yī gè) one/a/an, (c) 一匙 (yì chí) spoonful.

3.3.5.1.3. Word stress

The concept of stress for many is related to the clear pattern of discrepancy between stressed and unstressed syllables as found in English. Indeed, "stress" usually conjures notions of increased volume, increased duration, and elevated pitch while "unstressed" invokes notions of decreased volume, duration, and pitch as well as vowel reduction. However, many languages do not exhibit such a degree of difference between stressed and unstressed syllables.

3.3.5.1.4. Stress acoustic correlates

Fundamental Frequency, Frequency Range, duration, and intensity have been reported as the best prosodic acoustic correlates of stress. Not all of these are used in all languages and they are not all used in the same way to identify stress. For

instance, in English, stressed syllables show higher F0 (Fry 1955, 1958) but in Telugu stressed syllables have lower F0. (Balusu, 2001).

In the case of tone languages, the use of pitch – or frequency in Hz - for the production of tones eliminates, or lessens, the relevance of pitch as stress marker. Indeed, as increased pitch is the recognized marker of tone, if it is also used as a stress marker, it might be difficult to differentiate between unstressed and stressed syllables. This also entails that not stressing a syllable cannot be associated with the complete removal of F0 as this would lead to the "de-toning" of the syllable which, in turn, would lead to semantic ambiguity. Other phonetic features should thus be considered as markers of stress in tone languages.

3.3.5.1.5. Stress identification by native speakers

It would seem reasonable to postulate that the fewer as well as the smaller the differences exhibited by a language between stressed and unstressed syllables, the more difficult it would be for its speakers to differentiate stressed syllables from unstressed ones. This validates Duanmu's comment that:

> "In English, stress is usually obvious to the native judgment. In contrast, stress in Chinese seems much less obvious, and some linguists consider Chinese to have no stress [...]. Indeed, a lack of judgment on stress seems to be a property of tone languages in general." (Duanmu, 2004:9)

One must however consider that the native speakers' lack of intuition about stress placement does not necessarily mean that stress is non-existent. In fact, one should consider that some syllables in Chinese could bear the equivalent of an English primary stress and others the equivalent of an English secondary stress. The acoustic difference between these levels of stress has been little studied. Plag and Kunter's (2007) study did show however that the acoustic correlates of the distinction between primary and secondary stress are pitch, intensity, right jitter[28]

[28] Jitter is the high variation of glottal pulse timing; high jitter value is perceived as creaky voice. Plag et al. (2006) found that jitter is a significant predictor of the prominence relation in compounds.

and, to a lesser extent, duration. The authors also found that right prominence is not phonetically justified:

> Phonological prominence is phonetically only prominent in left positions, while with rightward phonological prominence, no syllable is phonetically prominent. Instead, the rightward pattern is acoustically characterized by more or less constant levels of pitch and intensity across the two positions. It may have been these phonetic properties of rightward prominence that made some scholars speak of "level" or "double" stress (Plag and Kunter, 2007)

In other terms, if, according to phonology, the rightmost syllable is identified as the stressed syllable, according to acoustic measures, both the left and the right syllables are equally stressed. Conversely, when the leftmost syllable is phonologically prominent (i.e. supposed to be stressed) acoustic measures confirm that it is also phonetically prominent (i.e. actually stressed).

Plag and Kunter (2007) also found that there is no acoustic justification of primary vs. secondary classification:

> A mixed effect regression analysis of the left position (with the vowel /ɔʊ̯/) reveals no significant effects whatsoever, which means that in the left position, primarily and secondarily stressed syllables do not differ in their stress-relevant acoustic properties. This clearly shows that stress is a relational property. [...] Whether the first prominent syllable is primarily or secondarily stressed is decided by the acoustic properties of the prominent syllable on the right.

It would thus seem that, in case of equivalent acoustic cues between two syllables of a word, listeners attribute primary stress to the right syllable, probably on the basis of an assumed word-internal declination. This could also be the reason for which Chao (1968) specified that two full-tone syllable words were stressed on the second element, the first element bearing medium stress and the second, heavy stress.

Duanmu (2004:9) acknowledges that if native speakers of SC cannot intuitively identify stressed syllables, they can easily identify toneless syllables such as the particles mentioned earlier. Stress in SC is also evident in words where one syllable is duplicated because, in such cases, the second syllable loses its tone. For example,

the word for "wife/Mrs" is the word "tài" pronounced twice as can be seen from its written form 太太[29]. However, the word is not pronounced with two 4th tones but with one 4th tone preceding a neutral tone.

Stress / tone relation: Duanmu (2005) postulates that, if one considers that the second syllable in such words is not just de-stressed but actually "de-toned", other issues have to be dealt with:

> One might argue that the difference between the word pairs [...] is not due to stress but due to the presence or absence of tone. If so, Chinese does not have lexical stress. This proposal has two shortcomings. First, it must explain why toneless syllables are phonetically unstressed [...]. Second, stress must be assumed for Chinese anyway, if not for word stress, at least for phrasal stress. So there is little theoretical gain in not recognizing word stress. (Duanmu 2005:19)

What the author identifies as two shortcomings have their own shortcomings. First, there is no *need* to explain *why* toneless syllables are phonetically unstressed. It might be interesting to verify that toneless syllables are always phonetically unstressed and that unstressed syllables are toneless – which will be discussed in the next part. It might also be interesting to attempt to explain why it is so; but there is no *need* to justify it to accept it. After all, we all accept that syllables undergoing vowel reduction in English are unstressed and nobody has ever argued that it would have to be justified - which would be more difficult as unstressed syllables do not all undergo vowel reduction. As for Duanmu's other identified inadequacy, suffice it to say that there is little theoretical gain in recognizing word stress if evidence against it were to be found.

[29] If the word were the addition of "tài" and its toneless "homophone" "tai", their sinograms would be different. For example "tài" (with the same tone) can also express among others "peptide", "attitude", "titanium", "extravagant", "to discard", "safe" but each of these words has a specific sinogram because a each sinogram has only one semantic meaning.

Comparison between English and SC: Duanmu (2005) argues that in fact the situations in English and in SC are the same. Part of the argument is that, in English, some words are made of two strong syllables between which the native speaker might find it difficult to identify the stressed one. However, this situation is not as common in English as it is in SC where these represent the large majority as illustrated in Table 2. Indeed, the difficulty in English of judging which of two strong syllables (column A) is stressed only occurs in a few expressions, possibly in a "modifier + noun" group and in a few words such as "downtown"[30]. In Chinese on the contrary, examples of those are plentiful. Furthermore, while the examples of clear stress placement (Column B) are abundant in English, in SC these are limited to cases of syllable repetition or cases of grammatical particles. Finally, there does not seem to be any examples of clear stress placement between two strong syllables (Column C) for SC while many are found in English.

Table 3.2: Comparison of disyllabic lexical items in English and SC. Syllables easily identified as stressed are underlined. No item with two strong syllables in which stress placement is intuitive was found for SC.

	A	B	C
	Unclear stress placement 2 strong syllables	**Clear stress placement** 1 strong syllable + 1 weak syllable	**Clear stress placement** 2 strong syllables
English	Red Cross green house (adj.+n.) downtown	Review, pardon	hotel, rabbi, greenhouse (compound)
Standard Chinese	mèi lì 魅力 (charm) lǎo shī 老师 (teacher)	mèi mei 妹妹 (younger sister) Dōngxi 东西 (thing / person)	

[30] The difficulty of stress placement on "downtown" is attested by the differences found in dictionaries. Indeed, this word is stressed, according to the *Oxford-Hachette French Dictionary*, on the first syllable, according to the *Longman Pronunciation Dictionary*, on the second syllable and, according to the Collins Cobuild, on both syllables.

From these data and analysis, it is apparent that stress and tone are inseparable in SC: stressed syllables have a tone while unstressed syllables are toneless and syllables with a tone are stressed while toneless syllables are unstressed. Does it mean that the correlates of stress are actually correlates of tone, and thus, that stress is part of tone in SC? With the information available at present, it is impossible to ascertain a positive or negative answer. Nevertheless, in this study, tone and stress will be assumed to be two different prosodic features, which happen to co-occur in SC.

Stress domain. As the existence of stress (associated with tone) is accepted here, the domain of stress must be established. We have seen that in English stress domain is the word while in French it is the phrase.

However, before we can evaluate if the word is the domain of stress in SC, we must define what a word is, a debatable issue in SC. As popular belief would have it, SC words are monosyllabic. This is in fact simply not true as thoroughly explained by, among others, DeFrancis (1984, 177-188) and Duanmu (2000:146-172). Partly responsible for the perseverance of this myth is the SC writing system. Indeed, the Chinese writing system, which requires a space between each sinogram – each of which is spoken as one syllable – does not help conceptualize polysyllabic lexical items as single semantic items.

According to Duanmu (2000), disyllabic items are a 20th century tendency and while monosyllabic words were formerly preferred, dissyllabic items are nowadays chosen. In fact, dissyllabic items are preferred even when the second item does not bring a nuance of meaning:

(4) méi 煤 = "coal"
tàn 炭 = "carbon/charcoal"
méi tàn 煤炭 = "coal"

(5) diàn 店 = "shop/store"
pù 铺 = "store"
diàn pù 店铺 = "store"

Some disyllabic items in fact illustrate the concept of disyllabic words in SC. Indeed, in some cases, the individual items, although associated with semantic value in the dictionary, are meaningless in themselves. DeFrancis (1984:182) gives the following example:

(6) pú 葡 = "grapes"
táo 萄 = "grapes"
pú tao 葡萄 = 'grapes'

The individual entries *pú* and *táo* carry the meaning "grape" but may not be used in a sentence without the other to mean "grapes". De Francis labels these syllables "CB: Completely bound".

According to He and Li (1987), of the 3 000 most commonly used SC words, 69.8% are disyllabic while only 27% are monosyllabic, 3% are of three syllables and 0.3 of four syllables.

As we saw earlier, it is impossible to determine that one syllable is more stressed than another and one should consider that both syllables are equally stressed. In a sense, this could be seen to reflect that SC speakers do not consider "mèi lì" a lexical item per se (with a specific semantic value) but still as the concatanation of *mèi* ("magic", "to charm") and *lì* ("power", "strength"). In other words, *mèi lì* would mean "the power to charm" rather than "charm". In contrast, disyllabic items which include a toneless or detoned second syllable (as is the case of *pú tao* 葡萄, above)

would correspond to the notion of a lexical item, a word, as generally perceived in non-tone languages. This claim finds support in the following observation by Peng et al:

> Containing a neutral tone syllable is one of the clearest indicators that a recurring sequence of syllables is a fully lexicalised polysyllabic word rather than a more decomposable compound or even a phrase, as in the monomorphemic Dōngxi 'thing' versus the obviously compound Dōngxī 'east-west'. (Peng et al., 2005:237)

What the authors call "fully lexicalised polysyllabic word" could be equated with long established compounds in English which are perceived as a lexical and semantic entity rather than as the connection between two original entities it is made of. For example, it could be argued that *blackboard* today conjures images of a childhood classroom while, in the eighteen hundreds, it would relate to an actual board that was black. Because *blackboard* evokes one specific meaning, it is stressed on one syllable only; in the noun group *black board* however, both the adjective and the noun are stressed.

Lin (2001:88) offers a list of pairs of two-syllable words (reproduced below in figure 3.35) the first one of which has a tone for each of the two syllables and the other only has stress on the first syllable. As the sinograms are the same for both pairs, there is little doubt that the two lexical entries shared, at some point, tones and meaning.

a.	i.	shízài	实在	'indeed'
	ii.	shízai	实在	'honest'
b.	i.	shēngqì	生气	'get angry'
	ii.	shēngqi	生气	'vitality'
c.	i.	liánzǐ	莲子	'lotus seed'
	ii.	liánzi	帘子	'screen'
d.	i.	duìtóu	对头	'correct'
	ii.	duìtou	对头	'rival; enemy'
e.	i.	dàyì	大意	'gist'
	ii.	dàyi	大意	'careless'
f.	i.	dìdào	地道	'underground tunnel'
	ii.	dìdao	地道	'genuine'
g.	i.	dōngxī	东西	'east and west'
	ii.	dōngxi	东西	'thing'
h.	i.	shìfēi	是非	'good and bad'
	ii.	shìfei	是非	'dispute'
i.	i.	biānpái	编排	'lay out'
	ii.	biānpai	编派	'fabricate'
j.	i.	xiōngdì	兄弟	'brothers'
	ii.	xiōngdi	兄弟	'younger brother'
k.	i.	sūnzǐ	孙子	'(name of a historical military strategist)'
	ii.	sūnzi	孙子	'fraternal grandson'

Figure 3.35: *List of pairs of two-syllable words in which the first word has a tone for both syllables and the second only has stress on the first syllable. Lin (2001:88)*

This author claims that, as only the second syllable of a two-syllable word can be unstressed, SC is left-stressed. However, if we retain the idea that these two-syllable words are compounds (in the sense that the semantic value of the item is not necessarily equivalent to sum of the two elements) then we can refute this proposal. Indeed, in English, compounds are stressed on the left-most element but this does not mean that all two-syllable items (be they two-syllable words or two monosyllabic words) are stressed on the first element.

In this sense, it could be asserted that SC is made of monosyllabic words (with tone) and of compounds (one tone syllable + one toneless syllable). The other cases of dissyllabic items will have to be considered concatenation of two words, as there is no evidence of those as single items at the prosodic level. In this perspective, a clearer graphic (pinyin) representation would to separate the syllables of the latter items but to attach the syllables of fully lexicalized items:

(7)a shí zài 实在 "indeed"

(7)b shízai 实在 "honest"

The conclusion of this section on stress is thus that "word" level stress in SC is limited to disyllabic items whose second syllable is inherently toneless (functional particles) or is "de-toned" because of tone duplication. In this sense, at this particular time, and at this phonological level, stress in SC is completely connected with tone:

There is stress if, and only if, there is tone.

Disyllabic items made of two strong (i.e. stressed and toned) syllables are pronounced with equal level of stress, even if "feels" like the second element is more stressed (see page 76).

3.3.5.1.6. Phrasal stress

Although no evidence of stress as an independent prosodic feature was found at the "word" level, a difference of degree of stress is in fact achievable in SC. Indeed, fluctuation of stress is used to show semantic contrast, i.e. to show the difference between narrow focus and broad focus. This contrastive stress is phonetically realized by an expansion of the pitch range (Jin 1996; Shen 1990:67; Cao et al., 2000:29). The actual direction of the expansion depends on the tone of the syllable

it is associated with: if the tone is falling, it will fall further when under contrastive stress and if the tone is rising, it will rise higher when contrastive.

As was seen with regard to the SC word, defining the SC "phrase" is no easy task:

> We note that in spoken Mandarin Chinese, instead of complete or complicated sentences, native speakers tend to speak in a sequence of phrases. These phrases, somewhat loosely governed by semantics [...] are grouped into perceptually identifiable larger units. Tseng (2003:599)

On the model of alphabetic languages, Chinese is often transcribed as strings of phrases separated by commas. For speech synthesizing, these phrases have generally been treated as sentences and have been given a sentential declarative intonation. However, the output gives the native listener an impression of unnatural abruptness and choppiness (Tseng, 2003:601). This feedback reinforces the assumption that phrase specific prosody does indeed exist in SC.

As prosodic segmentation is not intuitively obvious in SC, as well as because of a minimalist and unmethodical use of punctuation and an unclear relation between prosody and syntax, understanding the prosodic segmentation of SC requires a process which is the reverse of that used for other languages. Indeed, research into the prosody of non-tone languages was first guided by graphic and syntactic information: many researchers concentrated on those points which were visibly or intuitively recognized as boundaries to see if these junctures were reflected at the prosodic level. As graphic information is not helpful in suggesting SC prosodic segmentation, researchers now opt to look for prosodic boundary markers identified in non-tone languages. The assumption here is that prosodic markers found in non-tone languages will also be found in SC and will consequently provide information on SC segmentation. Among these prosodic markers, F0 reset, increased duration of preboundary segments, and breaks (i.e. pauses) have now been shown to be relevant in SC segmentation.

Also, the following rhythmic organization has been established (Cao et al, 2000): the prosodic word (PW), the prosodic phrase (PP) and the intonation phrase (IP).

The prosodic word (PW) is a segment of usually two or three syllables (although possibly of only one or more than three syllables). These segments, described earlier, are the lexical building blocks of (modern) SC and PW is the domain of tone sandhi and de-toning from tone duplication. Yang and Wang (2002:709) showed increased duration of PW-final syllables. These authors have also identified a slight pitch reset of the bottom line of intonation.

The Prosodic Phrase (PP) is "the most common and functional rhythm unit used in speech production and perception. Generally, it is larger than [a] word but smaller than [a] syntactically defined phrase or clause." (Cao et al., 2000:27) The same authors specify that this chunk of speech is usually limited to 9 syllables (7±2), which is similar to Tseng's (2002:3) 7.89 syllables found in what she calls the "major phrase boundary" (according to the "between B3 breaks" measure in her table, reproduced in figure 3.36 below). It also seems to correspond to the "clause" in Kochanski et al (2003:41): "We define [...] a clause as a subset of a sentence that is marked by a comma". According to (Cao et al., 2000:28), PP-final syllables are much longer than PP-initial syllables. Tseng and Lee (2004:253) confirm PP-final syllable lengthening and add that the third syllable from the end is shortened. Yang and Wang (2002:710) have also found that a significant pitch reset of the bottom line of intonation is typical of PP boundaries.

	M (Syllable # / ms)
B2	2.08 / 514
B3	7.89 / 1941
B4	13.15 / 3237

Figure3.36: *Tseng's (2002:3) between breaks number of syllables and length in ms.*

The Intonation Phrase (IP) corresponds to the syntactical sentence. Contrary to what was stated for PPs, IP-final syllables are shorter than IP-initial syllables (Tseng, 2002:28). In fact, Tseng and Lee (2004:254) found that the place of a PP in the breath group (i.e. in the IP) determines the lengthening or not of the final syllable. In other words, PP-final syllables are lengthened in IP-initial PPs but shortened in IP-final PPs. According to Yang and Wang (2002:710) a significant pitch reset of the bottom line of intonation and presence of silence is typical of IP boundaries.

In fact, segmentation of SC requires the identification of the type of boundary from which the type of prosodic unit preceding the boundary can be established. In many languages, this would be circular (the type of boundary indicates type of prosodic unit and the type of prosodic unit indicates the type of boundary.) However, in SC, as the type of prosodic unit cannot be identified by other means, it cannot be identified before the type of boundary is.

Yang and Wang (2002) summarize the types of breaks identified in SC are their respective acoustic correlates

> In conclusion, pre-boundary lengthening is the acoustic correlate of weak boundary. Pitch reset is that of medium boundary, while silence is that of strong boundary. Yang and Wang (2002:707)

No study has however reported higher or lower F0 association with PW, PP, or IP boundaries. For SC, we will thus need to use the system of boundary notation used by other authors.

B2 = prosodic word (PW) boundary

B3 = prosodic phrase (PP) boundary

B4 = Intonational phrase (IP) boundary

Pragmatic sentence level boundaries

Peng et al. (2005:248) identified two boundary tones: the High boundary tone and the Low boundary tone. These are aligned with sentence final (neutral tone) particles. This is reminiscent of Chao's (1968) Rising Ending and Falling Ending but the latter author described these to be associated with the voiced portion of the last syllable, toneless or not, and with anyone of the four tones. Peng et al illustrate their point with the sentence:

(8) tāmen bú mài yúsǎn ma?
3.PL NEG. buy umbrella interr. H%
Don't they sell umbrellas?
3.PL NEG. buy umbrella interr. L%
Well, but they don't sell umbrellas.

If the sentence-final particle *ma* is produced with a high boundary tone, the sentence means "don't they sell umbrellas?" with an element of surprise conveying that the speaker believed that the store would sell umbrellas. If the sentence-final particle *ma* is produced with a low boundary tone, the sentence is a statement – despite the sentence final interrogative particle – and means something like "well, but they don't sell umbrellas." According to the authors this prosody softens the speaker's explanation as to why he came back empty-handed.

3.3.5.1.7.Intonation[31]

To a great number of people, it would seem trivial to assume that native speakers of SC, like speakers of non-tone languages, are able to understand what type of sentence (declarative, interrogative, etc.) they hear and this even in the absence of

[31] Intonation is here understood as utterance boundary F0 variation (marked with % in ToBI tradition) that provides information significant for the whole utterance, irrespective of word and phrase level stress. In non-tone languages, intonation is normally associated with utterance final contours. (See 2.1.3)

syntactic markers such as *ma*. This assumption is however somewhat flawed as Yuan (2004) found that native speakers of SC occasionally misidentify interrogatives as statements.

In most, if not all, non-tone languages, intonation is marked by the use of a particular tone on the final or on the last stressed syllable of a sentence; a rising tone indicating an interrogative (devoid of syntactic marker) and a falling tone, a declarative. Because most syllables in SC already carry a tone, the application of intonation tones seems problematic. Various suggestions as to how intonation is indicated in SC have been put forward.

The first suggestion is that intonation is marked by a tone placed immediately after the last tone of the sentence. This "successive" tones approach was first proposed by Chao (1933). In this approach, a rising tone at the end of an interrogative will rise higher and for longer than that at the end of a declarative, while a falling tone in the interrogative context will simply give a succession of a falling (lexical) tone and a rising (intonational) tone. This suggestion entails that if a toneless syllable is situated at the end of the sentence, it will bear the rising or falling intonation tone.

This view is also adopted by Duanmu (2004) who explains that a boundary L can express a contextual meaning of affirmation and gives the following examples (2004:7):

(9)	H	+	L	→	H-L
	duo				*duo*
	many		affirm		"*certainly many*"
(10)	LH	+	L	→	LH-L
	nan				*nan*
	hard		affirm		"*Certainly hard*"

The author also clearly specifies that the tone succeeds rather than modifies the last lexical tone and that no extra syllable is introduced:

> The boundary tone L is realized after the original lexical tone is completed. Therefore the syllable is extra long, or equivalent to two syllables. Indeed, as Chao (1933: 130) points out, very often boundary tones are realized on a separate interjection particle in Chinese, such as [a]. (2004:7)

Duanmu also retains the idea of a boundary tone H used for interrogatives32 and provides the following two examples:

(11)	H	+	H	→	H-H
	duo				*duo*
	many		question		*'Many?'*
(12)	HL	+	H	→	HL-H
	man				*man*
	slow		question		*'Slow?'*

Other authors (e.g. Chang, 1958 and Ho, 1977) have put forward the idea of "simultaneous" intonation marking, meaning that the lexical tone and the intonation tone would be superimposed. In this perspective, the intonation tone influences the direction of the lexical tone. Thus, a rising intonation will emphasize a rising tone (e.g. tone 2), give a rising direction to the level tone (tone 1) and will thwart, at least partially, a falling tone (tone 4).

Besides these two suggestions (successive and simultaneous tones), in which intonation has a localized effect (at the end of the sentence), other propositions have also been put forward that specify a global effect. In these, it is understood that the intonation pattern has repercussions on the F0 of the entire sentence.

---

[32] Duanmu in fact writes that the boundary tone H can express the contextual meaning of a question. However, the term "question" is replaced here by "interrogative" because, as explained by Lee (2005:4), "As illustrated in the utterance, "Can you pass me the salt?" uttered at the dinner table, an interrogative sentence is not necessarily a question. »

One of these models is Gårding's grid model (1984, 1987), reproduced in figure 3.37 below. The grid refers to the space between two parallel lines in which lexical tones are realized. The direction of this grid can be level, rising or falling depending on the type of sentence (declarative, interrogative, etc.). The distance between the parallel lines, i.e. the range, is determined by the level of sentential stress. It has to be understood that, while this proposition shares the idea of a rising intonation for interrogatives, it entails the progressive rising of the entire sentence rather than a rising confined to the end of the sentence – be that the final syllable, or the final stressed syllable, or the final fully toned syllable.

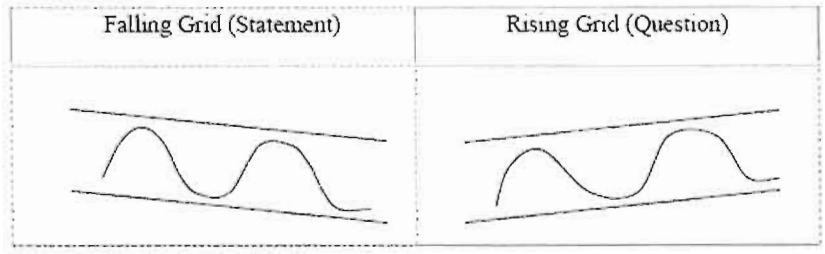


Figure 3.37 *Gårding's grid model (from Lee 2005:31)*

Shen, J. (1985) proposed a model similar to Gårding's but in which top and base lines are independent. The combination of directions of the two lines is used to differentiate between declarative, interrogative, imperative, and exclamatory sentence types.

Another global effect model is that of Shen X. (1990) who proposes three tunes. In her research, Tune I is used for statements and declaratives and is noticeable by the mid (frequency) starting point, mid-high highest point and low ending point. The other two tunes are used for interrogatives and both have a mid-high starting point and a high highest point. However, Tune II is used for yes-no questions (syntactically unmarked echo questions and syntactically "ma" marked questions) and exhibit a mid-high to high value ending point while Tune III is used for disjunctive questions (V + negation + V questions and alternatives) and is characterized by a low ending point – like Tune I.

These tunes are illustrated in Figure 3.38:

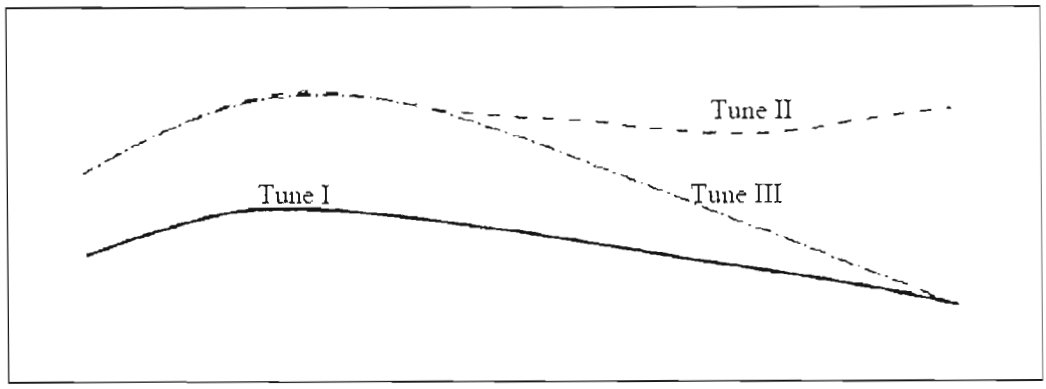


Figure 3.38: *Shen's Tunes.*[33] *Tune I is used for statements, Tune II for yes/no questions and Tune III for disjunctive questions.*

Two conclusions can be drawn from these findings: first, interrogatives start higher and have a higher highest point than declaratives do. Second, Tune II is equivalent to Tune I but raised a level or two.

The belief that interrogative intonation is raised is shared by many authors, among which De Francis (1963), Ho (1977), Shen (1985), Yuan et al (2002) and Zeng et al.(2004).

A more detailed study of yes/no questions in SC was conducted by Lee (2005). The author examined the intonation cues of two functions (information seeking and echo) of two types of yes/no questions (syntactically marked (*ma*) questions and syntactically unmarked questions). Lee finds that both types of questions display intonational cues. This means that syntactically marked questions are not spared intonation markers. However, markers are exaggerated in syntactically unmarked questions. The author also found that an expansion of pitch range is associated with a function, the echo function, rather than with a type of question. This range expansion is more salient over the sentence focus information (the part of the sentence subject in question) and even more so in echo questions. In addition Lee

[33] The legend of Shen's tunes has been amended here as the reference to the lines in her legend was incorrect with respect to her graph.

reports that yes/no questions also display more localized F0 rise and expansion of pitch across the last NP with greater F0 rise on the last high tone target (the last tone 1, 2, or 4 of the sentence). Finally, the study shows that, whatever the question type, echo questions are associated with moderately raised overall intonation and prominently raised top line over the last NP. On the other hand, information seeking questions are correlated with significantly raised overall intonation and pitch range becomes expanded over the last accentual domain.

Any conclusion on SC intonation has to start with the remark that no consensus on intonation cues and patterns has been reached. For speakers of non-tone languages, such an observation can be astounding as perception of intonation is, for them, intuitive and clear. Obviously, the use of frequency for both lexical tones and intonation accounts for the difficulty to judge which F0 variation is the realization of a syllabic and lexical tone pattern and which is that of sentential intonation contour. Some people have assumed that the pattern of the sentence contour comes first and that lexical tones graft themselves onto it, what Chao (1933) described as *"the small ripples riding on top of large waves"*. Others, like Hu (1985) and Wu (1982) have assumed that the overall intonation contour is created by the concatenation of lexical tones (i.e. that sentential prosody per se does not exist in SC). As we've seen, most authors now accept the idea of intonation in SC. This is substantiated by experimental studies that confirm that intonation contours vary according to their final tone and that the shape of a same tone varies according to the type of sentence it ends; the latter is demonstrated by Zeng et al. (2004) whose research show revealed that:

> "The tonal realisations were modified in a significant way according to the type of utterance. Tone 1, "flat", became ascending in rhetoric questions; Tone 4, descending", was characterised by a less accentuated slope; Tone 3, "descending-ascending", and Tone 2 which was realised in the same manner because of the contextual assimilation, weren't much affected by the modality, except that they possessed a higher register, and the second part of the contour (the reascending part) ascended much higher in rhetoric questions. (Zeng et al., 2004:3)

It remains to be seen whether intonation influences tones or tones influence intonation, or, as proposed by Xu (2004) with the PENTA model, the two work in parallel and with other communicative functions such as focus, demarcation, etc.

For the purpose of this work, we will accept the possibility of pitch raising (i.e. the use of a higher frequency throughout the interrogative sentence). This phenomenon is noted with *%q-raise*. We will also accept that an F0 rise can be seen near the end of a sentence to mark interrogatives; this rise can take place on the last syllable, on the last full toned or the last high tone target (tones 1, 2, 4) and will be marked with H%. Affirmative sentence declination will also be supposed to be similar to that of English and will be marked with L% and that the pitch range increases with the length of the sentence.

> *In Mandarin, it has been found that the rate of declination is faster at the beginning and slows down as the sentence progresses, which is compatible with what has been found in English. In addition, the overall F0 contour starts higher and ends lower in longer utterances (Shih 1997, 2000).* Lee (2005:18)

From the information gathered here, we can produce the following schema (figure 3.39), similar to Pierrehumbert's (1980:29) for North American English, which shows the possible tones and their association to form the grammar of intonation of SC.

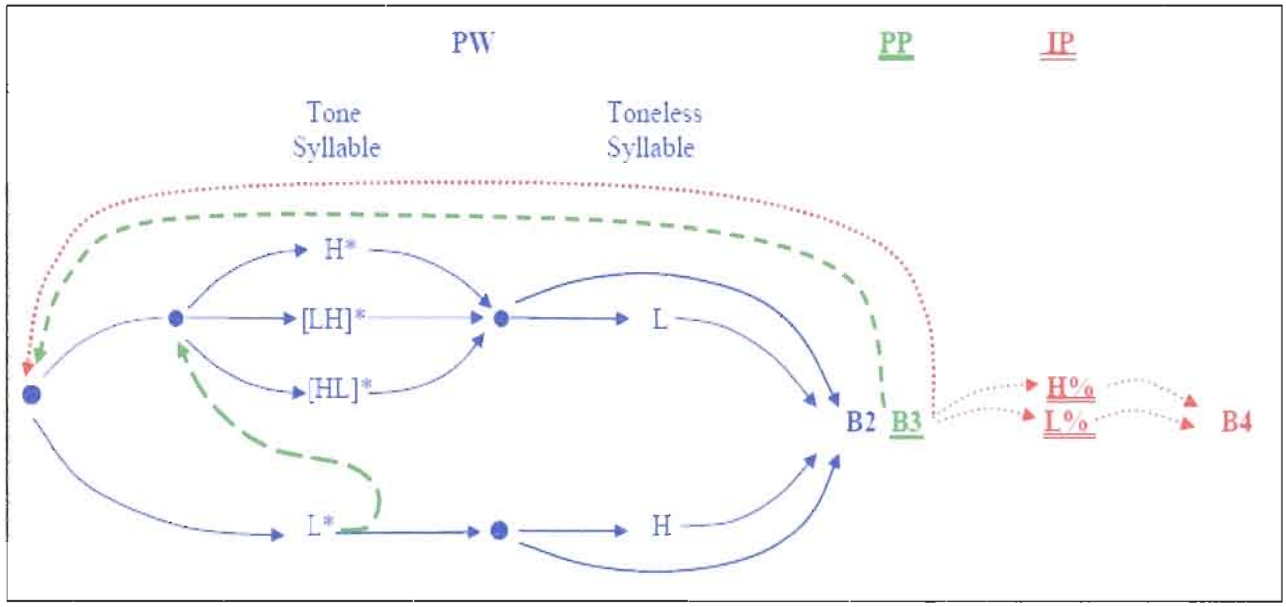


Figure 3.39 *Schemata of the possible tones and their association to form the grammar of SC. Note that the two-tone tones are bracketed to show that both tones are produced on the same syllable. Note also that the arrow after L* leads to the tones 1, 2, 4 node to avoid going back to tone 3 as it is impossible to have two consecutive tones 3 because of tone sandhi. It should also be noted that "toneless syllable" indicates the grammatical particles that are phonologically toneless but acquire a tone according to the previous tone; as these tones are associated with unstressed syllables, these L and H do not have a *.*

3.3.6. Discussion

In this chapter, the phonologies of North American English (NAE), of French (France and Quebec) and of Standard Chinese have been described independently. Time has now come to compare these phonologies because, if L1 phonology is transferred onto L2 phonology, this comparison will provide the information necessary to anticipate the prosodic errors made by learners of ESL. The comparison between the phonologies of NAE and the phonologies of the other two languages' will be done separately. After each of these, the errors one could expect native speakers of that language to make will be identified.

By comparing the grammars of NAE and of French as described in this research, the clearest differences between the phonologies of the two languages are:

1. North American English pitch accents are independent while those of French work in pairs.
2. The first pitch accent in the French AP is associated with the initial syllables of the AP while the last pitch accent in the French AP is associated with the final full syllable. English pitch accents are not associated with a syllable because of the syllable's location; pitch accents are however associated with the stressed syllables, the place of which is determined mainly by lexicon and grammar.
3. IP initial boundary tones are found in English but not in French
4. Phrase final accents are found in English but not in French
5. French AP final L is rare.

From the comparison of FF and QF, it was also found that:

6. The repetition of an LLH contour becomes an HLH contour for QF speakers.
7. Falling contours are associated with QF and flat contours with FF.
8. At the phonetic level, greater pitch range is used by QF speakers than by FF speakers. No comparative study of pitch range between the varieties of French and English is available.
9. FF speakers use a higher overall pitch than QF do. No comparative study of pitch height between the varieties of French and English is available.

We might thus anticipate the following phenomena in the production of English by native speakers of French:

1. Tendency to produce a pitch accent on the last syllable of all APs
2. Tendency to place stress incorrectly
3. The combination of the previous two points can lead to different situations if the stressed syllable is not the last syllable of the IP:

   a. Stress will not be produced on the normally-stressed-syllable but stress will be placed on the last syllable.
   b. Stress will be placed on the wrong syllable and on the last syllable
   c. Stress will be placed on the wrong syllable but not on the last syllable
   d. Stress will be placed on a correct syllable and on the last syllable
   e. Stress will be placed on the correct syllable.
4. Tendency, possibly more pronounced in native speakers of FF, to replace falling contours by flat or rising contours.
5. Tendency to produce an early pitch accent in all APs.
6. Non-production of phrase final accents. This could be more noticeable when a falling contour is needed as French AP final tones are rarely L*.
7. At least one of the two groups of native speakers of French is expected to use an inappropriate pitch range. That is to say, if English and QF use a similar range of pitch, FF speakers are expected to use an inadequately narrow pitch range. If, on the contrary, English and FF use a similar range of pitch, QF speakers are expected to use an inadequately wide pitch range. It is also possible that both varieties of French use an inadequately narrow pitch range or inadequately wide pitch range compared to English.
8. At least one of the two groups of native speakers of French is expected to use an inappropriate level of pitch. That is to say, if English and QF use a similar level of pitch, FF speakers are expected to use an inadequately high pitch level. If, on the contrary, English and FF use a similar range of pitch, QF speakers are expected to use an inadequately low pitch level. It is also possible that both varieties of French use an inadequately high pitch level or inadequately low pitch level compared to English.

By comparing the grammars of NAE and of SC as described in this research, the clearest differences between the phonologies of the two languages are:

1. The focal point of NAE grammar is the phrase (domain of the pitch accent) while in SC it is the word (domain of the tone)

2. In NAE, only stressed syllables may, but not necessarily do, bear a tone (called pitch accent). In SC, all stressed syllables bear a tone (lexical tone). Furthermore the vast majority of syllables in SC are stressed.
3. NAE tones not lexically relevant but SC tones are as the different tones on a syllable result in semantically different lexical items.
4. The most important prosodic markers are phrase accents and boundary tones in NAE and pauses in SC. Pauses are also found in NAE and boundary tones are found in SC but these have a lesser importance in terms of prosodic marking.
5. In NAE, part of a bitonal accent is aligned with the stressed syllable (an LH accent giving either L*H$^-$ or L$^-$H*). In SC, the accent is necessarily aligned with the syllable.
6. Pitch reset is characteristic of the boundaries in SC – slight reset at the PW level, significant reset at the PP and IP levels.
7. Pitch range, pitch contour and the place of the SC tone in the tonal space are very specific. For instance, tone 2 cannot be adequately described as LH because it goes up two levels only (on Chao's 1930 scale) and only from levels 3 to 5. In other words, production of LH from 2 to 5 or from 1 to 3 would feel wrong to native speakers. (Wang et al, 2003:5)

We might thus anticipate the following phenomena in the production of NAE by native speakers of Standard Chinese:

1. Tendency to give equal importance (stress) to each syllable. This can be expected to take the form of unreduced stressed syllables.
2. As tones in SC are attributed to stressed syllables only, if the SC speaker correctly produces an unstressed syllable, he/she might not attribute the H or L phrase accent.

3. Tendency of native speakers of SC to associate certain tones to certain syllables; for instance, a rising tone (tone 2) might be used to mark a stressed syllable.
4. Tendency of native speakers of SC to produce an L or H tone on unstressed syllables according to the preceding stressed syllables (as in a compound).
5. We might thus expect subjects to produce L*H^{-} and L^{-}H* in the same manner, that is to say, as [LH]*.
6. We might expect native speakers of SC to use pitch reset to show boundaries.
7. We might expect native speakers of SC to produce NAE accents as the specifics of SC tones. As Chao's (1930) letters do not correspond to F0, checking the production of specific tones in SC and the same speakers' production of pitch accent will be necessary.

In the next chapter, we will analyze the production of NAE by native speakers of French and the production of NAE by native speakers of SC to observe their placement and production of stress and the types of contours they used.

CHAPTER IV

EXPERIMENTS

4.1 Introduction

In this chapter, two groups of two experiments are presented. The first group of experiments is conducted with native speakers of French (QF and FF) and the second with native speakers of Chinese (SC). The experiments conducted with the two groups are different for a number of reasons. First, because points that are particularly salient in, and typical of, the L1 should clearly show if the L1 phonology is transferred. For this reason, the experiments examine, for French, early and late AP rises and, for Chinese, tone related phonological features. Second, because the SC speakers who participated in the experiments are students of the francophone University of Quebec in Montreal and therefore had to study French before they could study English. This could spoil the results obtained with SC speakers if they were to do the same experiments as those conducted on FF and QF speakers. Indeed, if the results were the same as those obtained with the Francophone subjects, it could not be determined whether this would be attributable to their L1 (Chinese) or their L2 (French). Finally, if the same experiments were conducted with both groups of speakers, one would have to assure that the judges were equally familiar with French accented and Chinese accented speech. Given the socio-political context of Quebec, no such assumption can be made.

While six judges were asked to independently classify all the speakers within the two language groups, different analysis methods were used for the different experiments, according to the needs. For instance, stress placement being easy to judge by native speakers of English, it was considered that one judge would be sufficient to analyse whether francophone speakers stressed long words adequately (experiment 2). As it happens and for reasons specified in 4.4.5, a single judge was not sufficient. In experiment 1, because of the subjectivity of having to judge if an

unwarranted stress was mild (presence of stress that does not distract too much from sentence rhythm) or severe (distracting the rhythm of the sentence), two judges were asked to individually listen to all sentences and a third judge was called upon to resolve any discrepancies between the first two judges' evaluation. For both experiments, when segmental features (for instance, unreduced vowel sounds) were suspected of influencing evaluation, the pitch contour of the segment was extracted to generate a pulse train (Praat's 'To hum...' command, Boersma and Weenink (2007).

For the two experiments in which English prosody produced by native speakers of Chinese is analyzed in terms of tone transfer and peak alignment, only instrumental analysis is used. Indeed native speaker judgement would be of no use in those cases.

Subject ranking and data collection took place in the same circumstances and at the same time for all the speakers and all the experiments. These are described hereafter.

4.2 Materials and procedures, stimuli, elicitation method

4.2.1 Materials and procedures

The materials included an Agreement to participate form, a short questionnaire, a short description of the nature and procedure of the experiment.

All forms and procedures (including the type of data collected as well as the means of collecting, storing, and using the data) were reviewed and approved by the "*Comité institutionnel d'éthique de la recherche avec des êtres humains*" of *Université du Québec à Montréal* (*UQAM*). All materials were written in French but English translation was made available.

The questionnaire provided information about the ESL subjects'

1. Name.
2. Telephone number.
3. Email address.
4. Native language(s) and variety of that native language (ex. France, Quebec, Haiti, Morocco, etc. for francophones and Beijing, Shanghai, Canton, Hong Kong, etc. for sinophones).
5. Foreign languages other than English.
6. Study at that session or at a previous session of one of six phonetics courses offered in the English programs at UQAM.
7. Authorization to access their administrative dossier.
8. Length of English studies.
9. Place of English studies.
10. Other experience that may have influenced their acquisition of English (for instance, immersion camp, private lessons, long period of time without studying and using English).

The procedures were explained to each participant individually to ensure that the subjects knew what to do and how to use Can8 Virtual Lab, software digital laboratory used to display the questions and to record the subjects' written and oral answers. They were also informed that they should produce the recordings as naturally as possible; "naturally" meaning "the way they would normally speak English, without making special pronunciation efforts". They were shown how to access and complete the different sections of the test.

All subjects were recorded individually in a computer lab using computers equipped with Creative Technology "Sound blaster live" sound cards and Sony HS-90B combined headphone and microphone headsets. All sentences were digitally recorded using Can8 Virtual Lab. The subjects were shown how to use the equipment and how to best adjust the volume of their recording.

4.2.2 Stimuli

All participants were asked to complete five groups of tasks.

The first part was a group of speech perception tasks in which perception of phonemic variation was tested. For example, upon hearing a recording, the subjects had to choose if they heard:

(1.) He made the putt yesterday

(2.) He made the pot yesterday

Another task included in this part was for subjects to say if the two sentences they heard were the same or different. No written form of the sentences was provided. For instance, upon hearing "They want fish all the time" and "They won't fish all the time", subject had to decide if they heard the same sentence twice or two different sentences and choose between:

1. Same sentence
2. Different sentences

This first part is not actually used in the following studies but enabled subjects to get accustomed to the system and to focus their attention on phonetic features.

The second part was a recording task. Subjects had to record the sentences they could read on the screen. A picture illustrating the sentence was provided to ensure the subjects understood the sentences. The elicitation sentences and pictures are provided in Appendix 1. Two target words were included in these sentences; each sentence had the word *fan* or the word or syllable *cat*. This is for potential comparison with the subjects' production of the same word/syllable in their native language, as explained in the description of part 5 below.

The third part was a speech perception task concentrated on perception of prosody. For instance, after being instructed that words written in green (here underlined) indicated increased stress, subjects were asked which sentence best represented what they heard.

Example:

He left her book on the black desk.
He left her book on the black desk.
He left her book on the black desk.
He left her book on the black desk.
He left her book on the black desk.
He left her book on the black desk.
He left her book on the black desk.

Subjects were also asked to judge if they heard *can* or *can't* in a few sentences – where phonemic information was not sufficient to judge between the two words. Some exercises included segmentation and the subjects were for instance asked to judge if they heard "she saw him cross the room" or "she saw him across the room". Another task required subjects to indicate how many syllables a word had and which of those syllables was stressed.

One of the tasks required subjects to indicate if they thought the sentence they heard was complete or not. For instance, they would hear

"This is typical of a ban"

and they had to decide if they had heard the complete sentence A or the incomplete sentence B (cut off after *aban*).

A. This is typical of a ban[.]

B. This is typical of aban[donment.]

This task tests subjects' ability to hear the difference between the falling contour (HL) that is used on the word "ban" and the high flat contour (H) that is used on the syllable *ban* of *abandonment*.

This part, along with native speakers' evaluation (see 4.3 below), was meant to be used to rank subjects in terms of prosody awareness (see 6.2.2 below).

The fourth part was another recording task. The sentences were more rhythmically complex than the sentences in part 2 and included polysyllabic words. The subjects were asked to record the six sentences in the order they were provided. More detail about the content of this section is provided in the appropriate studies.

The fifth and final part of the test invited subjects to record sentences in their native language. Francophones were given sentences that included the word or syllable /fæn/ as underlined in the following example:

Le fan club est formé d'enthousiastes mais pas de fanatiques.

The word "fan" and the first syllable of « fanatiques » could thus be used to compare the subjects' production of this syllable in English and in French. Other sentences included the syllable /kæt/.

Native speakers of Standard Chinese were asked to produce sentences that included /fæn/ pronounced with the four tones. Furthermore, each of these were placed in different places in sentences: (1) as the first syllable of a sentence, (2) mid sentence/phrase, (3) as the last syllable of an affirmative sentence, (4) before a comma, (5) as the last syllable of an interrogative sentence without the syntactic marker *ma*. Thus, native speakers of SC recorded 20 sentences in SC. All the sentences are presented in Appendix 2.

All participants were paid 5$ for their participation. They were told that this gratuity did not come with any obligation and that they were free to leave anytime they wished.

Elicitation method

The choice of using read speech rather than spontaneous speech is justified and explained by three main arguments.

First, this study consists of the comparison of prosodic elements in a same sentence or phrase produced by different speakers and the comparison of specific items (ex. *fan*) in different languages produced by the same speakers. Moreover, the sentential position of the target words needed to be controlled. For these three reasons, conversational speech, even carefully elicited, would not have provided the necessary data.

Second, the competence[34] level of these participants is generally not high enough for them to have developed different registers of pronunciation. Indeed, the participants, mainly ESL students, speak in the classroom the way they speak in the real life. For instance, their pronunciation errors are the same when they participate in class and when they talk to each other or to their teacher in informal contexts, recess or encounters outside class. In fact, their greatest, if not sole, use of English is in the classroom.

Finally, emotions and attitude of the speaker and toward the listener affect prosody (see chapter 2). Read speech has the advantage of providing emotion free data while conversation speech characteristically includes some form of affect. This is particularly important in this research as individuals of different nationalities react differently to the presence of a interviewer. For instance, Quebeckers feel quite at ease with professors whom they address using the casual "tu" form while students of France are much more inhibited and could not consider using anything but the polite/formal "vous" form. As the speech of these two groups is compared, read speech eliminates the emotional issues and places all subjects on an equal emotional footing.

[34] In this thesis, the term "competence" refers to both "competence" and "performance" as defined in generative linguistics.

This choice of read speech reflects the needs for this particular work. It is however undisputable that for other purposes and in other circumstances conversational speech would be preferable.

4.3 Ranking of the subjects

Two methods were originally used to rank subjects' level of competence in prosody. The first method was to measure subjects' ability in perception of prosody; this was done with part 3 of the test described above.

The second method was to have six judges perceptually evaluate each subject's production of prosody. These judges are all native speakers of NAE or at native speaker level of competence in NAE. Four of the judges have training in phonetics; two had little or no training before the experiment. All were given the same instructions to consider only prosodic features such as stress placement, stress production (pitch pattern, etc.), presence/absence of pause, length of pauses, rate of speech, fluency/smoothness (staccato, choppy, etc.), linking. They were told to ignore segmental aspects (choice of vowel or consonant, quality of vowel sound or consonant sound, etc.)

The judges were asked to listen to all four sentences presented as listening to only part of the sentences might not offer a complete prosodic picture. The four sentences were taken from Part 2 and included two shorter ones and two longer ones. For ease of use, the data were presented in the format of a multiple choice exercise, for which the possible answers were the following levels of competence:

1) Beginner
2) Elementary (low)
3) Elementary (mid)
4) Elementary/intermediate
5) Intermediate (mid)
6) Intermediate (high)
7) Advanced
8) High advanced/fluent

What each judge considered each level to mean did not matter as ranking, not evaluation, was needed. In other words, the importance lies in the level of each subject compared to the others. What "elementary/intermediate" means, in absolute term or with regard to a program of studies, does not matter. All that matters is that each judge understood that "elementary/intermediate" is higher than "elementary (mid)" and lower than "intermediate (mid)". Furthermore, whether some judges might be more familiar to Francophone ESL speech than Sinophone ESL speech did not matter as only intra (not inter) language group ranking was relevant.

The judges' choices were then converted into points; 1 point for beginner, 2 points for low elementary, and so on. The judges' individual and combined assessment for all three groups of speakers (FF, QF, SC) are provided in Appendix 3. The judges' assessment was consistent.

No correlation between competence in perception and performance in production of prosody was supposed. For instance, it was not presumed that perception is acquired before production or vice versa. The only possible assumption was that the subjects who ranked higher in the perception tasks might rank higher in the production task.

No correlation was however found between the subjects' level of perception and level of production. In fact, the results provided by the perception test were not judged to be reliable because some obviously advanced subjects scored lower than

obviously elementary subjects. One of the possible reasons for these results is that some subjects had the advantage of having followed speech perception courses. Indeed, the material for both the experiment and the courses were created by the same person and students who took the speech perception courses were familiar with some of the material and the format. As the vast majority of students had received pronunciation training at some stage in their studies and as this training was not necessarily provided by the person who did the experiment, the impact of pronunciation training was not considered to be responsible for the discrepancy between the subjects' level in speech perception and in pronunciation. Furthermore, as the experiments aim at identifying evidence of phonological transfer in the production (not the speech perception) of ESL, the pronunciation ranking was used in the experiment and the perception results were disregarded.

4.4 Study 1: French

4.4.1 Purpose of study

The purpose of this study is to analyze the production of NAE by native speakers of French (QF and FF). The areas to be analyzed are some of the potential difficulties brought forward from the comparison of the grammars of NAE and of French. The specific objectives of each experiment will be explained in the appropriate section.

4.4.2 Subjects

Five native speakers of NAE, all teachers of ESL, provided the control recordings. The details about the speakers of French are provided in the tables below[35].

[35] More details about individual speakers of all groups are provided in Appendix 6.

Table 4.1 Native speakers of FF

Number of subjects:	**12**
Number of male subjects:	6
Number of female subjects:	6
Average age:	**3,3**
Youngest subject:	23
Oldest subject:	40
Average years of ESL learning:	**3**

Table 4.2 Native speakers of QF

Number of subjects:	**21**
Number of male subjects:	6
Number of female subjects:	15
Average age:	**35**
Youngest subject:	23
Oldest subject:	61
Average years of ESL learning:	**7**

More than half the students reported that they were resuming their study of English after a long period without using English.

Most participants were studying ESL; a few participants were studying another subject and took no or very few English courses; three participants were university employees.

In the following experiments, the subjects' ranking (lowest to highest in the subject's language group) is born in mind to see correlation between the feature examined and the subjects' level of proficiency.

4.4.3 Experiment 1

4.4.3.1 Objectives:

The objective of experiment 1 is to investigate the production of rhythm[36] by native speakers of French in simple English sentences.

4.4.3.2 Procedure and Stimuli

Procedure: The speech of native speakers of FF and QF is analyzed separately and compared later.

Stimuli: The sentences that the subjects read are short and include only monosyllabic words. These sentences were also simple in their syntax and lexicon (to avoid hesitation); moreover, a picture was provided to illustrate the sentence.

The three sentences used in the following analyses are:

Sentence 1: The fan is blue and white.

Sentence 2: The first fan on the left is red.

Sentence 3: The black cat plays in the snow.

4.4.3.3 Hypotheses and Analysis

Hypothesis 1: The rhythm of sentence 1 will be similar whether the segmentation is done with NAE phonology or French phonology. Because of this, one expects native speakers of NAE and francophone subjects to apply similar rhythm to this sentence.

For instance, in the sentence "The fan is blue and white" the words "fan", "blue" and "white" are expected to be stressed and the others unstressed. All stressed words would be associated[37] with a F0 peak:

---

[36] Rhythm is here defined as all prosodic features (F0 valleys and peaks, intensity rises and falls, non-intrinsic segmental length variation, absence/presence of pauses and length variation of pauses) except intonation. See chapter 2.

[37] To say that peaks are (phonologically) associated with certain syllables does not suggest that they are (phonetically) aligned with them.

The **fan** is **blue** and **white**

%L T* L- T* T* L- L%

All pitch accents are indicated with *. However, no specific pitch accent tone has been written as a few of them (H*, L*, etc.) are possible; instead the letter T (tone) is used to mark the presence of a tone. The phrasal accent L- can be expected to mark the end of the two phrases ("the fan" and "is blue and white"). As the sentence is a declarative, the default L% marks final boundary tone. If using French phonology, the segmentation of the sentence is as follows:

[The fan] [is blue and white]

[L H*] [L Hi L H*] L%

We thus see that the two segmentations predict the words "fan", "blue", and "white" to be stressed. The actual tones associated with each of these is only predictable for French (H*).

Hypothesis 2: Francophone speakers will segment Sentence 2 appropriately, that is to say, into three phrases (APs): The first fan / on the left / is red.

However, from the phonological segmentation, we can expect some difficulty from FF and QF speakers in the production of "on the left". Indeed, if the Hi is produced, stress will be associated with either "on" or possibly "the", both of which should not be stressed. For the same reason, FF speakers might attribute stress to "is" of the last AP.

[The first fan] [on the left] [is red].

[L Hi (L)H*] [**Hi** L H*] [**Hi** H*] L%

The two incorrect Hs (in bold) would be Ls in native speakers' segmentation.

Hypothesis 3: Non native speakers might segment sentence 3 into three APs, therefore creating the situation as explained for Sentence 2, in which stress is attributed to a normally unstressed syllable, in this case, "in".

[The black cat] [plays] [in the snow].

[L Hi (L)H*] [Hi] [**Hi** L H*] L%

The incorrect H (in bold) would be an L in native speakers' segmentation.

Hypothesis 4: For all three sentences, it is expected that native speakers of French of France will use rising pitch accents at the end of phrases, unlike Quebec French and NAE speakers who are likely to use falling pitch accents.

Hypothesis 5: Articulation rate is expected to positively correlate with subjects' ranking. Indeed, articulation rate has been shown to reflect the level of fluency of L2 speakers as well as those speakers' level of proficiency as evaluated by native listeners (Trouvain and Gut, 2007). Lower ranked subjects are also expected to use more and longer pauses. This will further contribute to a slow speech rate for these speakers.

Hypothesis 6: All difficulties will occur more frequently and more severely (longer pauses, greater stress, etc.) in the speech of lower proficiency level subjects.

Analysis

Each speaker's production of the three sentences was perceptually evaluated for adequate production of stress on each word by two judges who individually listened to all sentences and by a third judge who was called upon to resolve any discrepancies between the first two judges' evaluation. The judges had to evaluate if the words that should be stressed were sufficiently stressed and if words that should not be stressed were indeed unstressed. The judges were also asked to evaluate if an unwarranted stress was mild (presence of stress that does not distract

too much from sentence rhythm) or severe (distracting the rhythm of the sentence). When segmental features (for instance, unreduced vowel sounds) were suspected of influencing evaluation, the pitch contour of the segment was extracted to generate a pulse train (Praat's 'To hum...' command, Boersma and Weenink (2007).
Each speaker's production of the three sentences was also analyzed with Praat for contour of pitch accents at the end of phrases (combined with perceptual judgment), presence of pauses and length of pauses, and sentence length.

4.4.3.4 Results[38]

Sentence 1:

Stress: As was expected, all subjects produced sentence 1 (The fan is blue and white) appropriately. For instance, no stress placement mistake was recorded: the words "fan", "blue", "white" were stressed[39] while "the", "is" and "and" were not.

Contour: All phrase final pitch accent contours were within normal range. That is to say that none of the contours produced by FF subjects were rising so sharply or so high that they sounded inappropriate.

Pauses: some of the lower proficiency level QF subjects inserted pauses in one or both of the intervals (Int) specified here: *The fan* (Int1) *is blue* (Int2) *and white.* Int1 pauses are unsurprising and was even produced by one of the native speakers. Int2 pauses are however unusual. Pauses in both these intervals were produced more often by lower level subjects and were also longer when produced by lower level subjects.

[38] The quality of the sound files of two speakers for these sentences was quite poor. The recordings of speakers FF08 and of QF28 were therefore removed from this experiment.
[39] While the words "the", "is" and "and" were adequately unstressed, some subjects did not reduce vowels as should be done in English. Vowel reduction is however beyond the scope of this work.

Only three FF speakers produced an Int1 pause. Two of these were similar in length (41ms and 62ms) to the one produced by the native speaker (52 ms). The third pause was so long (250ms) that it indicates that the speaker probably encountered a problem (distraction, difficulty reading the stimulus) more than it indicates a phonological pause. None of the FF speakers produced an Int2 pause. No francophone speaker produced pauses in other intervals.

Figure 4.1 recapitulates the findings on pauses. The table regroups first, native speakers, second, FF speakers and finally QF speakers. The subjects of the last two groups are organized in the ranking order (lowest to highest) established by the judges. Larger scale figures are available in Appendix 5.

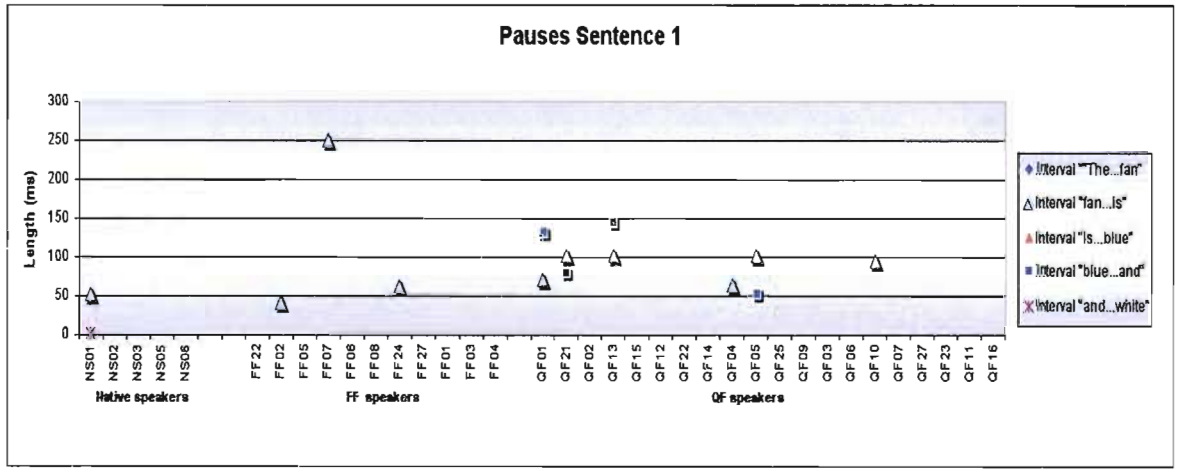


Figure 4.1: *Occurrence and length of interval pauses in sentence 1 produced by individual speakers at the different boundaries. Speakers are presented in order of rank (ascending) within their linguistic group (native speakers, FF speakers and QF speakers). Large- scale figure available in Appendix 5.*

Sentence 2:

Stress: As was expected, some subjects inadequately produced sentence 2 ("The first fan on the left is red"). Indeed, while all words that should be stressed (*first*, *fan*, *left* and *red*) were stressed, some of the words that should not be stressed, were. This is the case of "on" and "is". As these are the first word of a phrase (respectively, "on

the left" and "is red."), to native speakers of French they should be associated with an early AP rises (Hi).

This inadequate stressing was found in the speech of lower level subjects. Indeed, all FF speakers and the lowest ranked first 10 speakers of QF all stressed one or both these words. Stress errors are summarized in figure 4.2 below. Each incorrect use of stress was perceptually judged as either moderate (the shorter cones) or more significant (the higher cones). The cases of more significant stress are produced by the lowest level groups of both FF and QF speakers.

Two cases of excessive stress on "the" in "on the left" were also detected. These cases were preceded by significant stress of "on".

No excessive stress was found on "the", first word of the sentence. No de-stressing of the content words was found. Relative stress between "first" and "fan" is discussed in the next section.

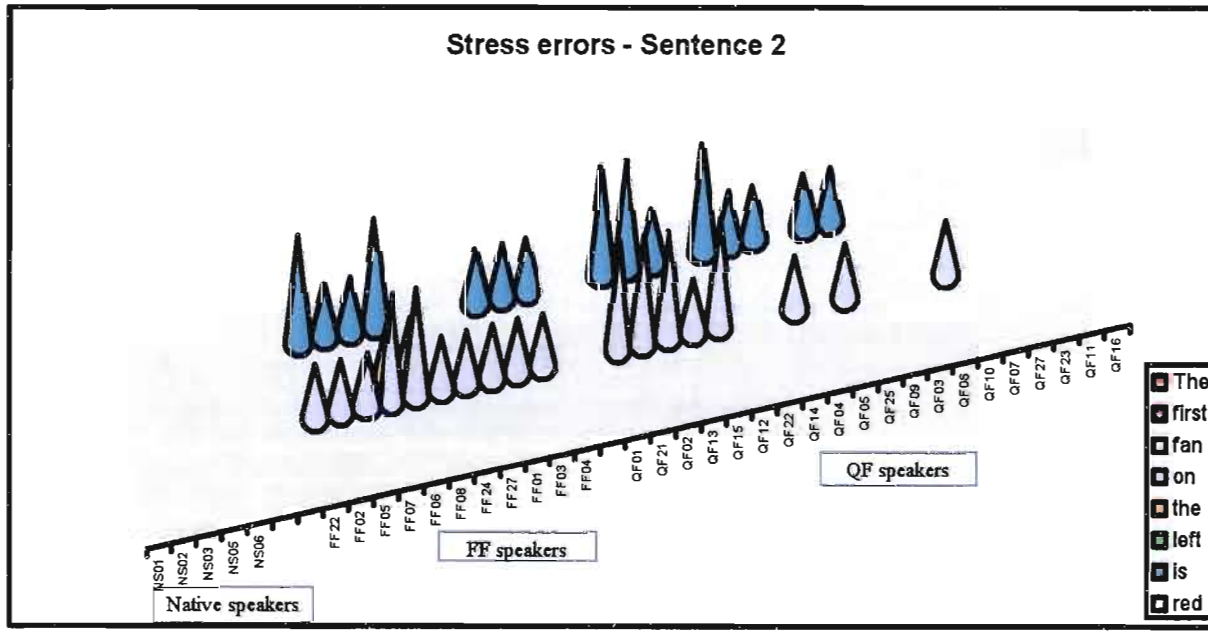


Figure 4.2: *Stress errors produced by non native speakers. Cones indicate errors, that is to say, stress of a function word. Smaller cones represent moderate stress; taller cones represent more significant stress. Large- scale figure available in Appendix 5.*

Contour:

The pitch contour used by the five native speakers, on the last syllable of "the first fan" is a falling or a falling rising contour.

Only two of the FF speakers produced such a contour while the others produced a rising contour except for two who produced a flat tone. Those who produced an appropriately falling tone were those ranked higher. The effect of rise in three of the speakers is emphasized by the lower stress used in the word "first". Indeed, typically, the ordinate "first" will bear the focus of the phrase (4.3). Three of the (lower level) subjects produced that word at a lower level than "fan", while three subjects produced the two words with similar stress.

Four QF speakers produced a rising tone on "fan" and two others produced a flat tone. No correlation between this and the subjects' ranking can be made as they are evenly distributed in the first two-thirds of the ranking. Only one QF speaker, who also produced a rising tone, produced "first" with less stress than fan. Two speakers, ranked at one third and two thirds of the range, produced "first" and "fan" with equal stress.

We can therefore say that, as posited by the hypothesis, a greater proportion of the FF speakers (4 out 11) than of QF speakers (4 out of 20) produced a rising tone or a flat tone on the phrase final syllable, as shown in table 3.

Table 4.3: Comparison of FF speakers and QF speakers' production of pitch contour on the word "fan".

	Total speakers	Rising tone	Flat tone	Falling tone
FF	11	4	2	2
QF	20	4	2	11

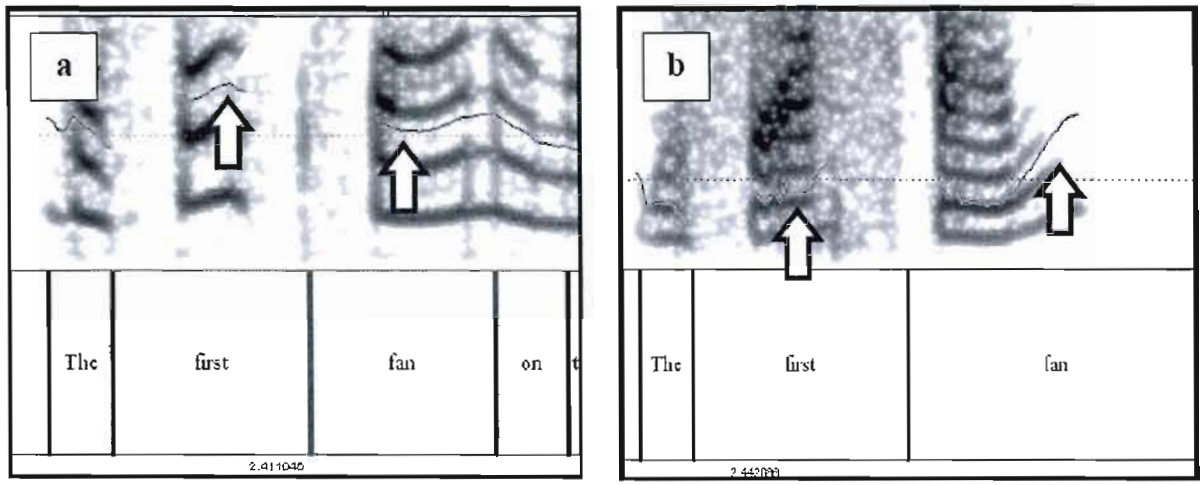


Figure 4.3: *Typical production of "the first fan" by a) native speakers and b) by L2 speakers. In a), the ordinate "first" bears the focus of the phrase while in b), "fan" bears the focus.*

Pauses: some of the lower level QF subjects inserted pauses in one or both of the intervals (Int) specified here: *The first fan* (Int1) *on the left (*Int2) *is red.* As phrase boundaries, both intervals may receive pause. None of the native speakers produced a pause in either interval.

Int2 pauses are however unusual. Pauses in both these intervals were produced more significantly by lower level subjects and were also longer when produced by lower level subjects.

Of the eleven FF speakers, five produced an Int1 pause and three produced an Int2 pause. Of these, two speakers produced both pauses. The production of both pauses and the production of longer pauses are found among the lower ranked subjects. However, not all low rank subjects produced pauses.

Three of the lower ranked speakers and two higher level speakers produced both pauses. Mid range, four speakers produced the Int2 pause only. No clear correlation can be made between rank and production and length of pauses though, as some low

rank speakers did not produce pauses. The lowest ranked subject also produced an inappropriate pause between "the first" and "fan".

Figure 4.4 recapitulates the findings on pauses.

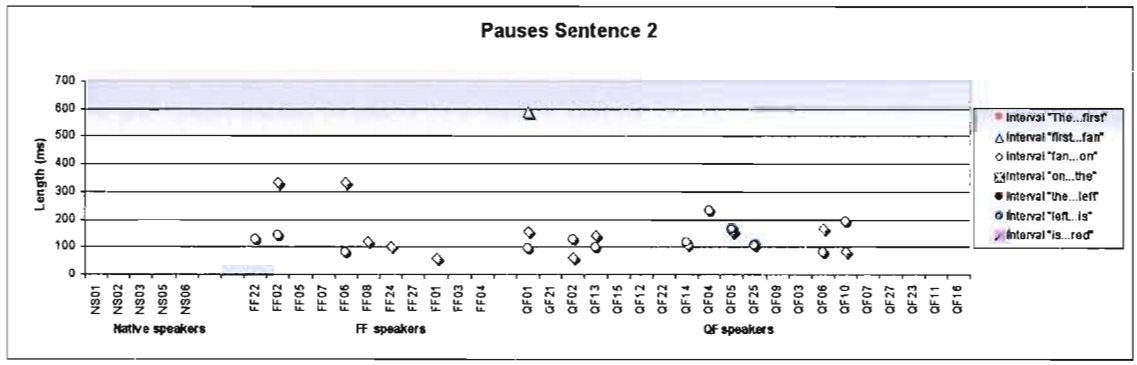


Figure 4.4: *Occurrence and length of interval pauses in sentence 2 produced by individual speakers at the different boundaries. Speakers are presented in order of rank within their linguistic group (native speakers, FF speakers and QF speakers). Large- scale figure available in Appendix 5.*

Sentence 3:

Stress: As was expected, some subjects inadequately produced sentence 3 ("The black cat plays in the snow"). Indeed, while all words that should be stressed ("black", "cat", "plays" and "snow") were stressed, some of the words that should not have been stressed, were. This is the case for "in" and, to a lesser extent, for "the". As "in" is the first word of a phrase ("in the snow"), to native speakers of French it equates to early AP rises (Hi).

The severity of inadequate stressing is correlated with subject ranking. Although the lowest ranked FF speakers did not stress "in", the next seven FF speakers did. Of these, the lowest ranked also stressed "the" and the first four produced more significant quantity of stress.

The situation is very similar among QF speakers as twelve of the fourteen lowest ranked speakers produced superfluous stress. Among these, the four lowest ranked speakers were judged to produce more significant quantity of stress. A speaker of

mid range rank stressed "the", probably because of hesitation. All stress errors produced by FF and QF speakers in sentence 3 are presented in figure 4.5.

No excessive stress was found on "the", first word of the sentence. Relative stress between "black" and "cat" is discussed in the next section.

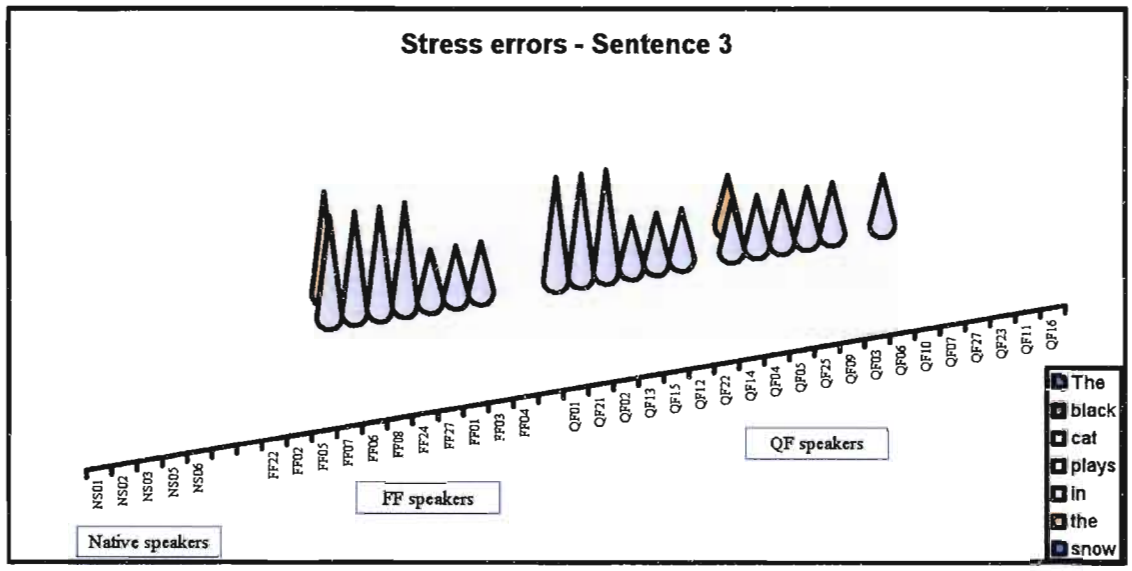


Figure 4.5: *Stress errors produced by non native speakers. Cones indicate errors, that is to say, stress of a function word. Smaller cones represent moderate stress; taller cones represent more significant stress. Large- scale figure available in Appendix 5.*

Contour:

The pitch contour used by the five native speakers on the last syllable of "the black cat" is a falling contour.

Seven FF speakers produced such a contour, one produced a rising contour and three a flat tone. The three lowest ranked subjects produced a flat or a rising tone.

Five QF speakers produced a rising tone on "cat", six produced a flat tone and ten produced a falling one. Although three of the rising tones were produced by very low rank subjects, no correlation can be found between rank and pitch contour as a rising contour was also produced by the second best subject and a mid ranked subject. All native speakers stress "cat" more strongly than its adjective "black".

Five FF speakers produced this same pattern, two produced "black" more strongly, and five gave both words equal stress. No correlation between rank and stress distribution can be found.

Only one QF speaker, who also produced a rising tone, produced "first" with less stress than fan. Two speakers, ranked at one third and two thirds of the range, produced "first" and "fan" with equal stress. Twelve QF speakers appropriately stressed "cat" more than "black", one produced "black" more strongly than "cat", and eight gave both words equal stress. As for FF speakers, no correlation between rank and stress distribution can be found.

We can therefore say that, contrary to the hypothesis, fewer FF speakers than QF speakers produced a rising tone on the phrase final syllable as summarized in the following table:

Table 4.4 Comparison of FF speakers and QF speakers' production of pitch contour on the word "cat".

	Total speakers	Rising tone	Flat tone	Falling tone
FF	11	1	3	7
QF	20	5	6	10

Pauses:

Three of the native speakers produced a pause between "black cat" and "plays" (Int1) but none produced one between "plays" and "in the snow" (Int2). The other two native speakers did not produce any pauses. The three lowest ranked FF speakers produced both interval pauses. Only one other subject produced an Int1, probably due to hesitation or distraction as the excessive length (434ms) indicates.

The lowest ranked half of the QF speakers produced one or both interval pauses. The length of these pauses decreases as the subjects' rank increases. Atypically, one speaker ranked quite high also produced the two interval pauses. Figure 4.6 recapitulates the findings on pauses.

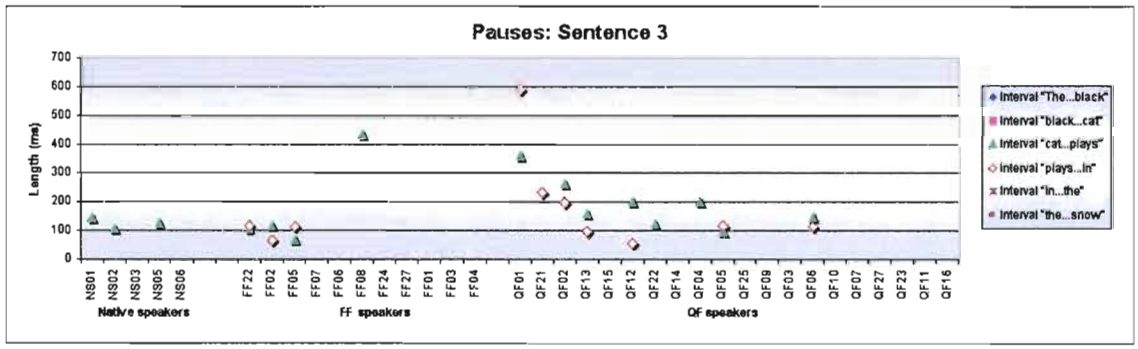


Figure 4.6: *Occurrence and length of interval pauses in sentence 3 produced by individual speakers at the different boundaries. The first five speakers are the (unranked) native speakers. The rest of the speakers are presented in order of rank within their linguistic group (FF speakers and QF speakers). Large- scale figure available in Appendix 5.*

Rate of speech:

As was hypothesized, L2 speaker ranking positively correlates with rate of speech. Indeed, the results indicate that the lower the student's level, the slower the rate of speech (figure 4.7). This is noticeable even if the pauses are subtracted from the overall sentence length (figure 4.8).

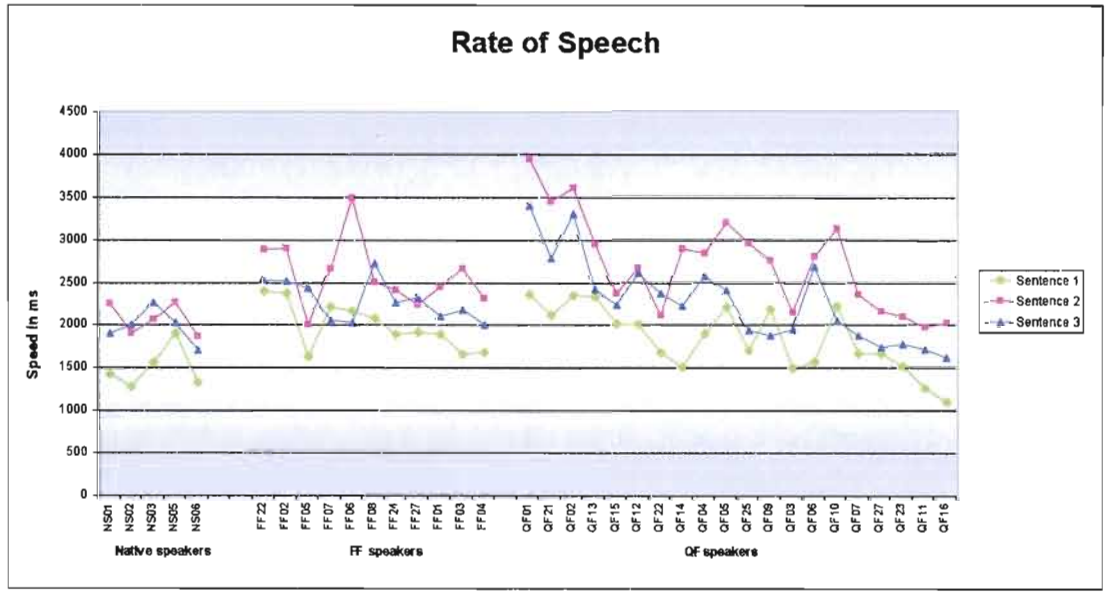


Figure 4.7: *rate of speech in ms per sentence per speaker. The first five speakers are the (unranked) native speakers. The rest of the speakers are organized in their groups (native speakers, FF speakers and QF speakers) in order of rank. Large- scale figures available in Appendix 5.*

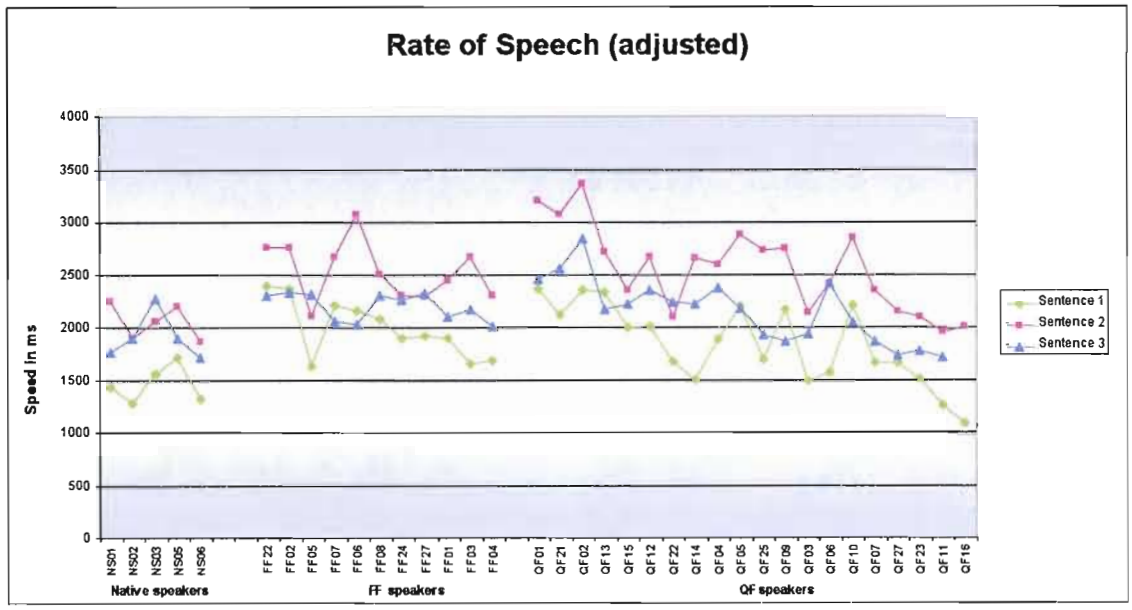


Figure 4.8: *rate of speech in ms from which pauses were removed. Speakers are organized in their groups (native speakers, FF speakers and QF speakers) in order of rank. Large- scale figures available in Appendix 5.*

4.4.3.5 Conclusion

In all three sentences, all content words were adequately stressed. Function words were correctly unstressed except for those that were at the beginning of a word group, for example, prepositions that introduced a prepositional phrase. In these cases, the function word was interpreted as early AP segments and was stressed, adequately with regard to French phonology but inadequately for English phonology. Figure 4.9 illustrates the combined stress errors in sentences 1 and 2. It also shows that the higher concentrations of these errors were made by lower level speakers of both FF and QF; similarly, the more severe errors (represented with taller cones) were made by the lower ranked speakers among those.

The first hypothesis predicted that francophone subjects would generally produce adequate rhythm for the three sentences. The second and third hypotheses specified that the function words "on" or possibly "the" (sentence 1) and "in" (sentence 2) would be stress. The results support those hypotheses.

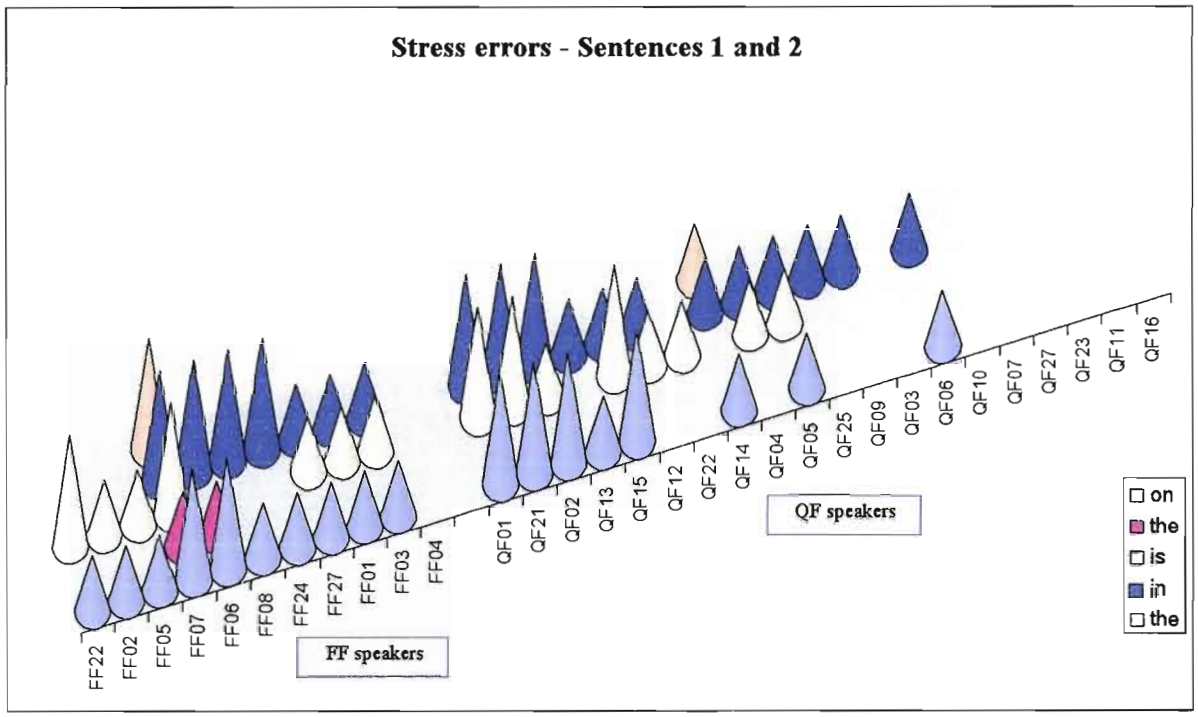


Figure 4.9: *Combined francophone ESL speakers' stress errors in sentences 1 and 2 for function words.*

The fourth hypothesis, which predicted that FF speakers would use rising pitch accents where QF and native speakers would not, was only partially confirmed. Indeed, while a higher proportion of FF speakers produced a rising contour on "fan" in the first sentence, a higher proportion of QF speakers produced such a rise in the second sentence. Furthermore, no correlation between rising contours and speaker rank could be found.

Clear correlation between articulation rate and subject rank was however found, as posited by hypothesis 5. The suggestion that more pauses and longer pauses would add to the slow rate of speech of lower ranked subjects was also confirmed.

We can therefore say that hypothesis 6, which stipulates that all difficulties would occur more frequently in the speech of lower proficiency level subjects, was substantiated except for the use of rising contours (hypothesis 4).

4.4.4 Experiment 2: Production of rhythmically complex sentences.

4.4.4.1 Objectives

The objective of experiment 2 is to investigate the production of rhythm by native speakers of French in rhythmically complex English sentences. Both the production of polysyllabic words and the general rhythm of the sentences will be examined.

4.4.4.2 Procedure and stimuli

Procedure: Subjects' production of the rhythmically complex sentences was analyzed using both auditory-perceptual and instrumental techniques.

Auditory-perceptual analysis:

Two judges individually listened to all sentences and marked all syllables they perceived to be stressed. A third judge was called upon to resolve any discrepancies between the first two judges' evaluation. When segmental features (for instance, unreduced vowel sounds) were suspected of influencing evaluation, the pitch contour of the segment was extracted to generate a pulse train (Praat's 'To hum...' command, Boersma and Weenink (2007).

Instrumental analysis

Fundamental frequency, duration, and intensity, recognized as the best prosodic acoustic correlates of English stress, were assessed in the subjects' production of the sentences with the Praat software.

Stimuli: The sentences that subjects recorded were rhythmically more complex than those used in experiment 1 and included polysyllabic words. These sentences were however syntactically and semantically simple.

Sentences analyzed:

Prison offered him the perfect opportunity to study.

The carpenter is impatient and motivated.

4.4.4.3 Hypothesis and Analysis

English stress is often called "free". This is because, unlike in other languages, stress placement varies from word to word. English stress is however not free in the sense that speakers are at liberty to stress the syllable of their choice. For a lexical item with a grammatical function, one specific syllable may be stressed. English stress placement must therefore be learnt. Typically though, in the western world at least, it is not taught alongside vocabulary in classes other than pronunciation classes. Indeed, students learn the graphic representation and the semantic value of lexical items first, and pronunciation later. Furthermore, phonemes are considered central to pronunciation teaching while stress placement and stress placement rules are deemed peripheral. (See 2.4 for more detail).

French phonology, on the other hand, requires the placement of a pitch accent on the last syllable of all APs, even when the AP is a single word. Lexical items are all marked to potentially bear stress on the last syllable.

Lexical stress

We can thus anticipate that, when confronted with the production of English polysyllabic words not normally stressed on the last syllable, native speakers of French will adopt one of the following strategies:

1. Stress will not be produced on the normally-stressed-syllable but stress will be placed on the last syllable.
2. Stress will be placed on the wrong syllable as well as on the last syllable
3. Stress will be placed on the wrong syllable but not on the last syllable
4. Stress will be placed on a correct syllable as well as on the last syllable
5. Stress will be placed on the correct syllable only.

We might further anticipate that the most erroneous stress placement (proposition 1) would be produced by lower level subjects who are not conscious of English prosody at all. It is considered most erroneous here because it is simply a transfer of French phonology. Propositions 2 and 3 are likely to be produced by subjects of a level higher than beginner. This would show that they are aware of English prosody being different from that of their language but not familiar with the actual rules, or that they are mistaken about stress placement in that word. The last two hypotheses would be expected of high level subjects. In both propositions 2 and 3, speakers would show that they are aware that NAE stress placement is different from French stress placement but ignore which syllable should be stressed. The difference between the two propositions however is that, in proposition 2, speakers would concomitantly produce (erroneously) NAE stress placement and (correctly) French stress placement. Similarly, in both propositions 4 and 5, subjects would obey English stress placement, but in proposition 4 speakers also respect French stress placement.

Sentence rhythm

Looking at the rhythm of the carrier sentences from the phonological perspective rather than the lexical perspective, we might presume different segmentations according to subjects' competence ranking.

As low level students tend to concentrate on one word at a time and pronounce each word separately, they might segment the first sentence as follows:

[Prison] [offered] [him] [the] [perfect] [opportunity] [to] [study]

We can further imagine that they would interpret each segment as a complete thought group (i.e. for those subjects, an AP). This means that in the case of the two-syllable words of the sentence, which should all be stressed on the first syllable, subjects would either stress the second syllable of the word or stress both syllables. Indeed, the native stress pattern on these words (H* L) is particularly rare in French as the early L and late H are least likely to be dropped out of LHiLH* (see 3.4.4.2.1). Clearly this result is the same as the one obtained in the lexical stress analysis.

With the segmentation presented above, one should note that "him", "the", and "to" could also be interpreted as full APs and thus be stressed. However, as native speakers of French are familiar with determiners and morphological markers/prepositions, they are likely to associate them with "perfect" and "study" respectively. The segmentation would therefore be:

[Prison] [offered] [him] [the perfect] [opportunity] [to study]

As subjects' competence increases, the segmentation is likely to evolve in the following manner:

[Prison] [offered him] [the perfect] [opportunity] [to study]

[Prison offered him] [the perfect opportunity] [to study]

In all these segmentations, "him" is the last syllable of an AP. We can therefore expect francophone subjects to stress this word, which is not stressed by native speakers.

As both the determiner "the" and the morphological marker "to" are the first syllable of an AP, they are likely to be assigned the L of the early rise and thus not be stressed, which is adequate here.

The segmentation of the second sentence should not create problems because the sentence has the same grammatical structure as the easiest sentence evaluated in Experiment 1 (see 4.3.1.3.3).

Experiment 1: *The fan is blue and white.*

Experiment 2: *The carpenter is impatient and motivated.*

The subjects' production of these sentences will therefore be analyzed for stress placement and production in polysyllabic words.

4.4.4.4 Results

Hereafter are the results of the judges' auditory-perceptual analysis of stress placement and rhythm in the subjects' production. Each word of the two sentences (below) is first discussed individually and then in thought groups.

Prison offered him the perfect opportunity to study.

The carpenter is impatient and motivated.

1. The word "prison" was appropriately produced by the vast majority (90%) of subjects. Only the two lowest ranked QF subjects stressed both syllables equally. One mid ranked subject produced some stress on the second syllable although it was not judged to be stressed as strongly as the first syllable was. No subject stressed the second syllable only. These findings are summarized in figure 4.10 below.

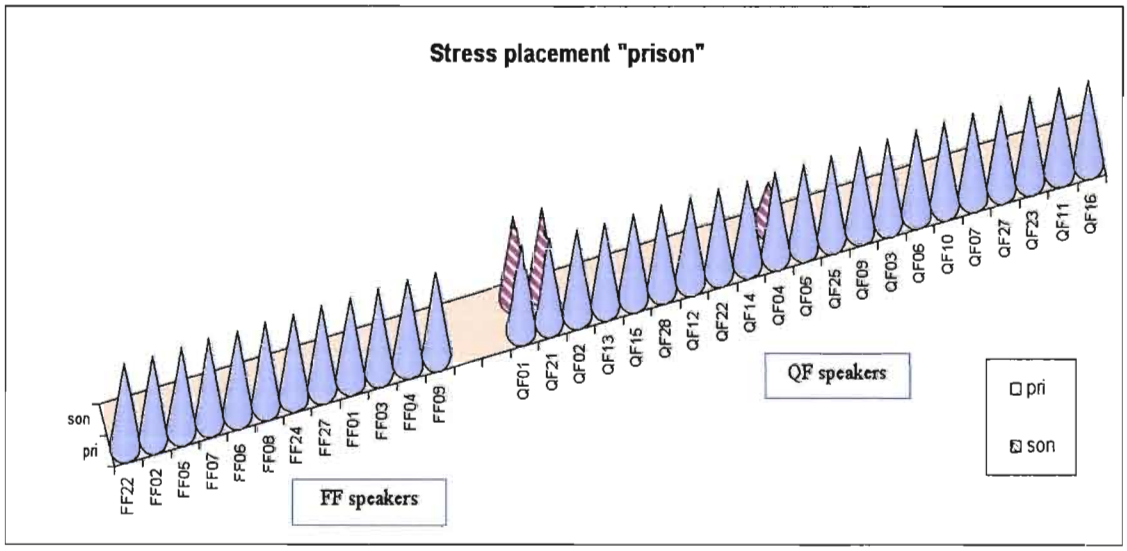


Figure 4.10: *Production of stress in "prison" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

2. The word "offered" was not appropriately produced by the majority of subjects.

Indeed, 15 out of 33 subjects stressed the second syllable of this word.

The three lowest ranked FF speakers stressed only the second syllable.

Two other FF speakers stressed the second syllable as well as the first.

Three QF speakers stressed the second syllable only. These ranked 9th, 17th and 18th out of the 21 QF speakers.

Seven of the 11 lowest ranked subjects produced stress on both syllables.

These findings are summarized in figure 4.11 below.

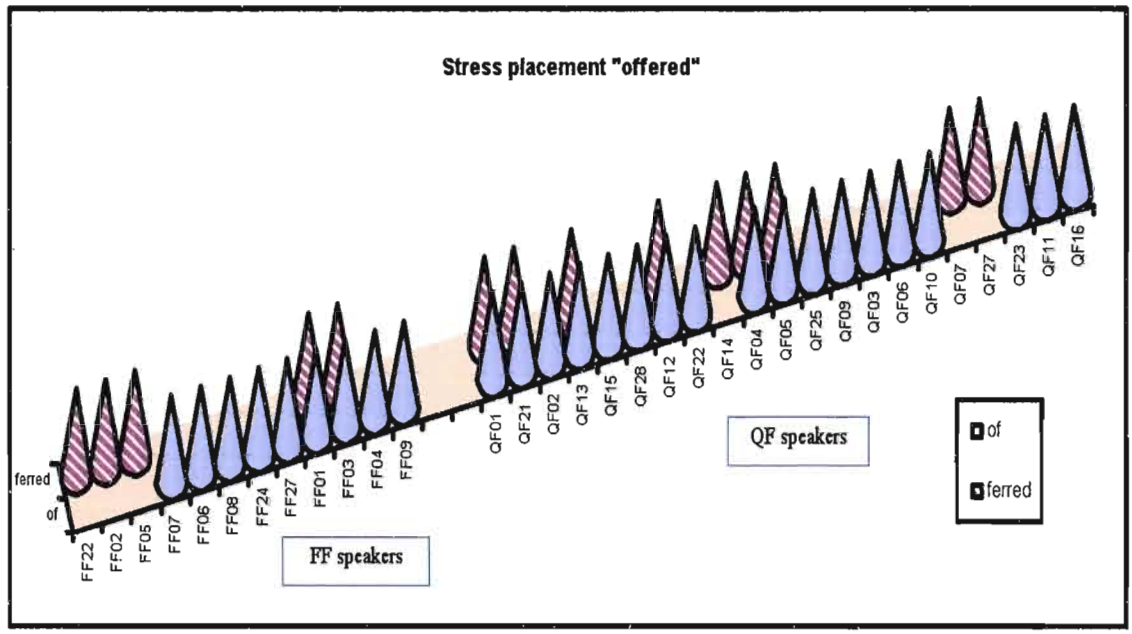


Figure 4.11: *Production of stress in "offered" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

3. The pronoun "him", AP final syllable, was stressed by all FF and all but the five highest ranked QF speakers. Some of these stresses were considered to be milder than the others but were stressed enough to sound unnatural to native speakers. Stressing the AP final syllable thus seems to be rank related for QF speakers. As the FF speakers' overall level was similar to that of the lowest 13 QF speakers (see appendix 3), no FF speaker does not disprove the finding that AP final stress is rank related as it is possible that a FF speaker of high enough level would not have stressed "him".

These findings are summarized in figure 4.12 below.

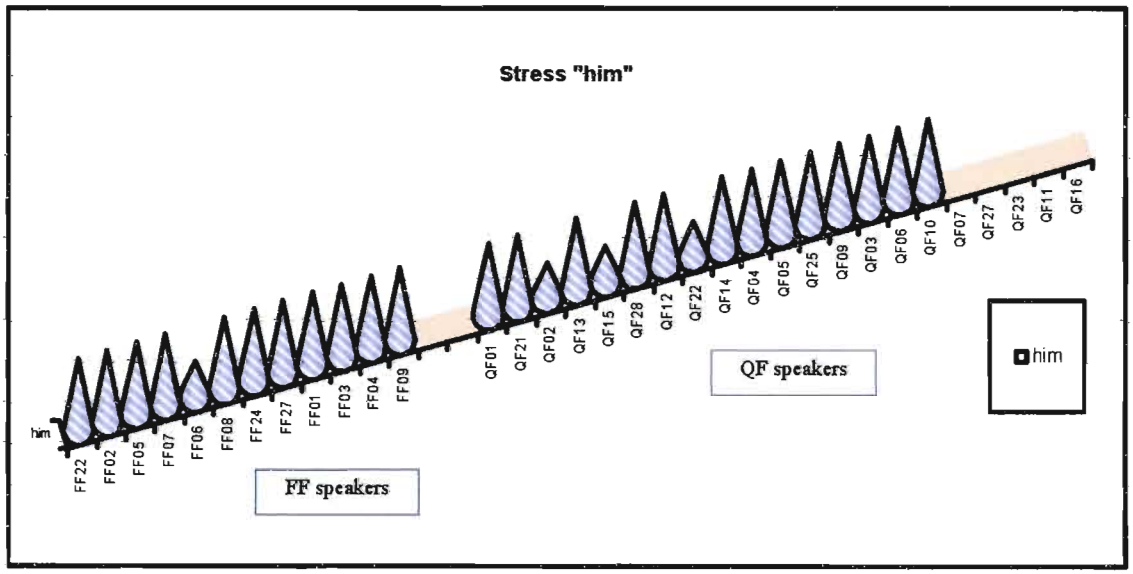


Figure 4.12: *Production of stress in "him" by FF and QF speakers. All cones signify a stress error. Large- scale figures available in Appendix 5.*

3. None of the speakers stressed the determiner "the"
4. The adjective "perfect" was correctly produced by all but two lower graded QF subjects.
5. These findings are summarized in figure 4.13 below.

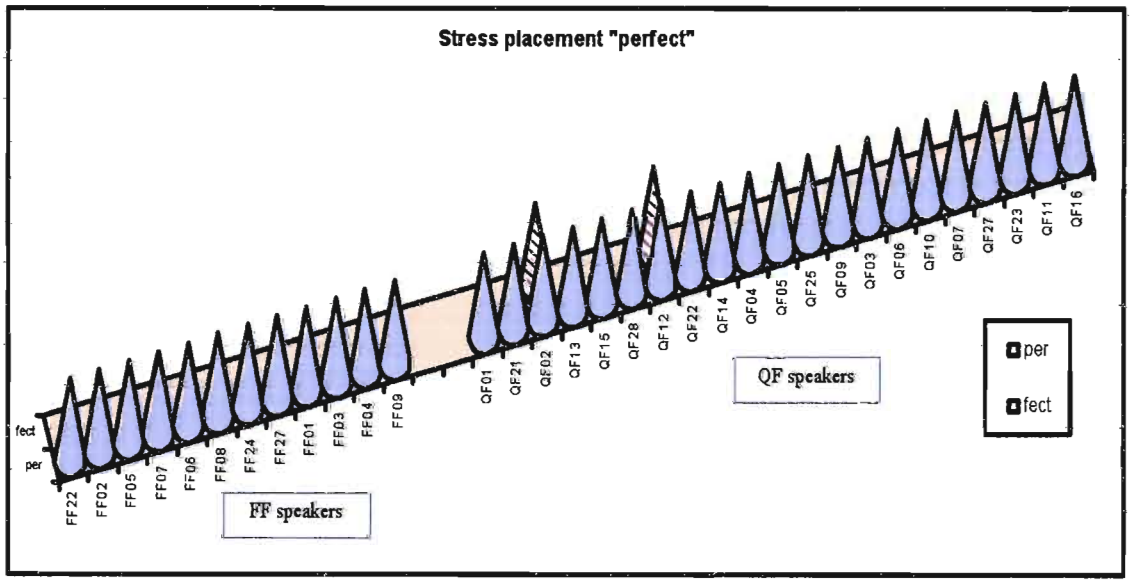


Figure 4.13: *Production of stress in "perfect" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

6. The word "opportunity" was correctly produced by the majority of subjects. None of the speakers misplaced the stress as all of them produced the syllable "tu" with stress. Some speakers however also stressed either the syllable "op" or "por". As the syllable "op" normally bears secondary stress[40], if it was marked as being stressed, it means that it was too stressed compared to what a native speaker would produce. In other words, it means that in these subjects' production, the syllable "op" was perceived to be "as stressed as" or "more stressed than" the syllable "tu".

A stressed "op" could thus be a sign of incorrect relative stress production while a stressed "por" indicates a stress placement error.

These findings are summarized in figure 4.14 below.

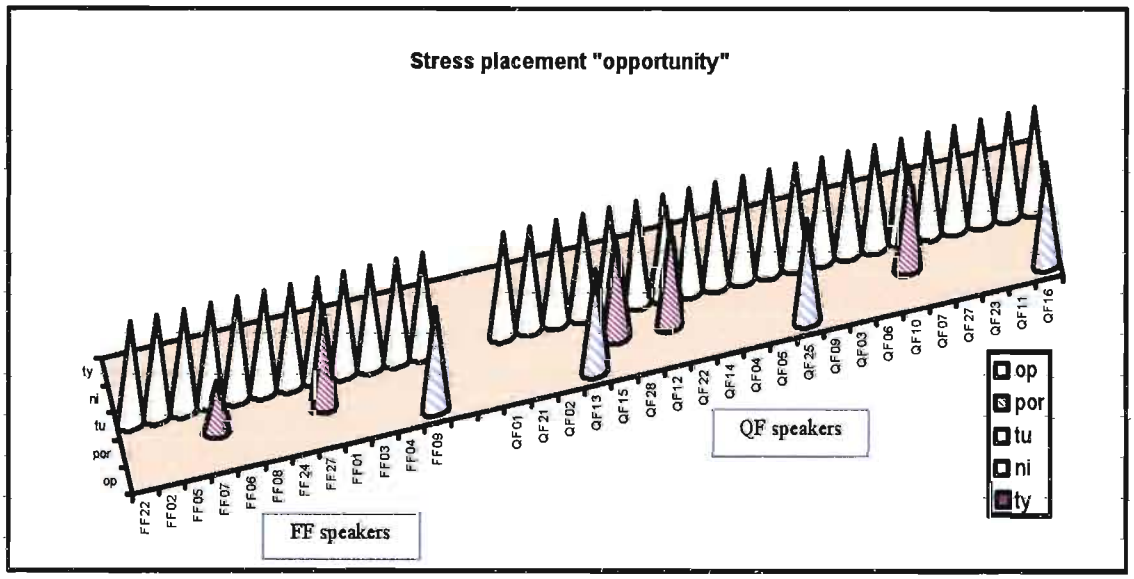


Figure 4.14: *Production of stress in "opportunity" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 2.*

[40] Referring to Kunter and Plag's 2007 research (see Phonology of Standard Chinese) both "op" and "tu" are equally stressed but the latter is perceived as more stressed. As the subjects' production is perceptually analyzed, the term "primary" and "secondary" stress are highly relevant and will be used here.

7. None of the speakers stressed the morphological marker "to"
8. The verb "study" was correctly produced by the majority of students. Only two FF subjects stressed the second syllable as well as the first. These findings are summarized in figure 4.15 below.

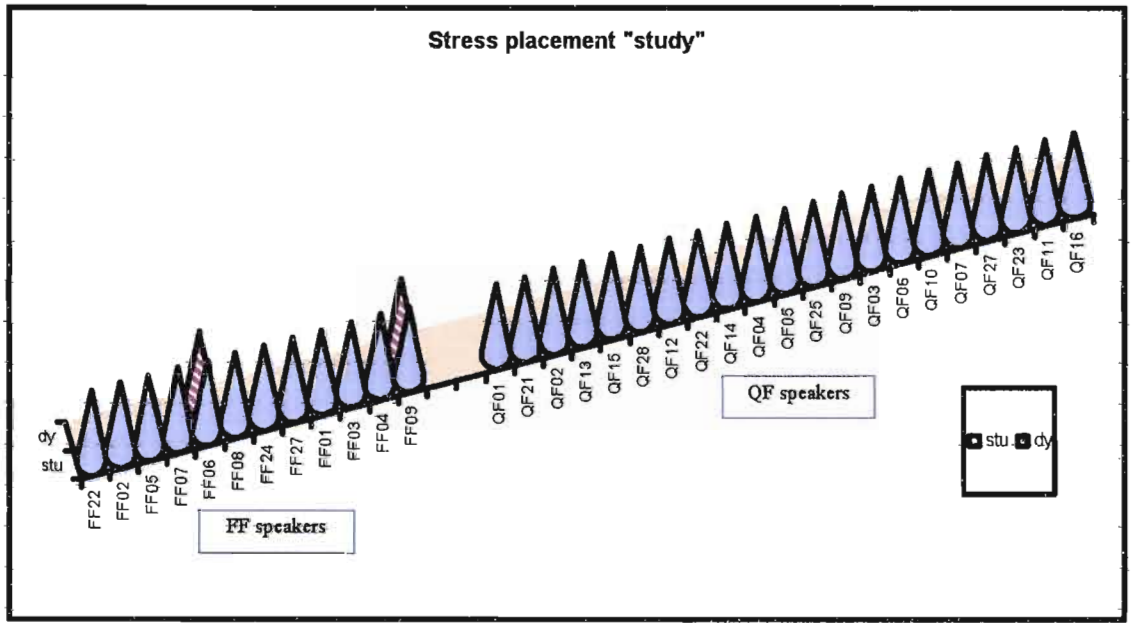


Figure 4.15: *Production of stress in "study" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

In the second sentence, the word "carpenter", normally stressed on the first syllable, was correctly pronounced by only 11 of the 33 francophone speakers.

Of the 12 FF speakers, 2 stressed the first syllable, 4 the second syllable only and 6 stressed two syllables (the first syllable and the second or the first syllable and the third).

Of the 21QF speakers, 10 stressed the first syllable, 2 the second syllable only and 6 stressed two syllables (including the first syllable), 1 the last two syllables and 2 stressed all three syllables.

No correlation can be made between stress placement and subject ranking. These findings are summarized in figure 4.16 below.

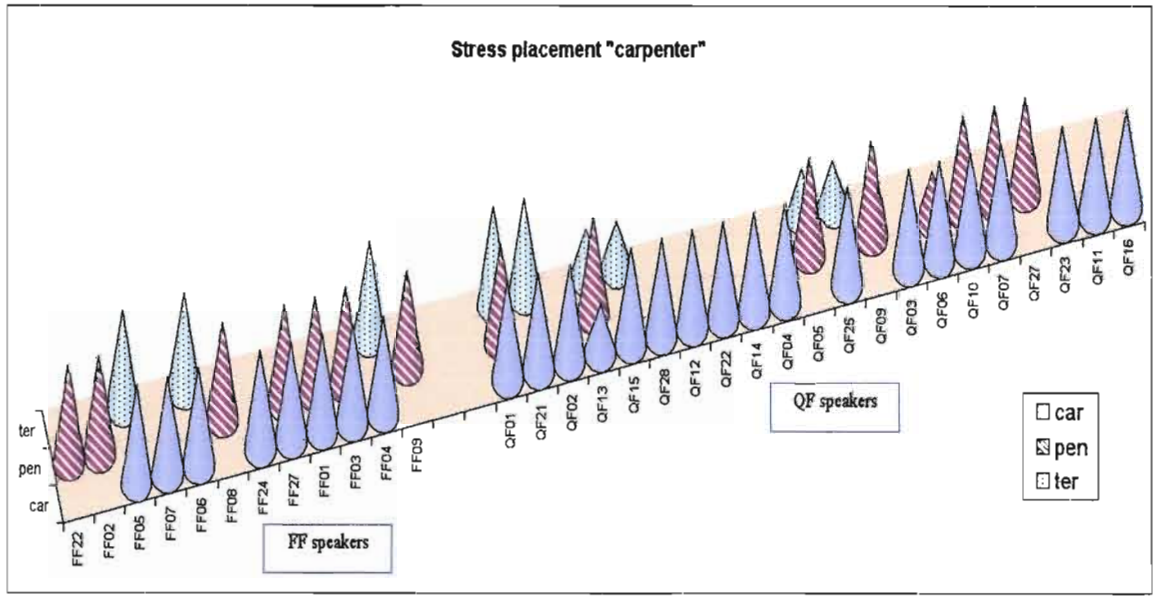


Figure 4.16: *Production of stress in "carpenter" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

The word "impatient" was adequately stressed by 9 of the 12 FF speakers and by 15 of the 21 QF speakers. The other three FF speakers stressed the first syllable. Among the remaining 6 QF speakers, 4 stressed the first syllable, and two stressed all three syllables.

These findings are summarized in figure 4.17 below.

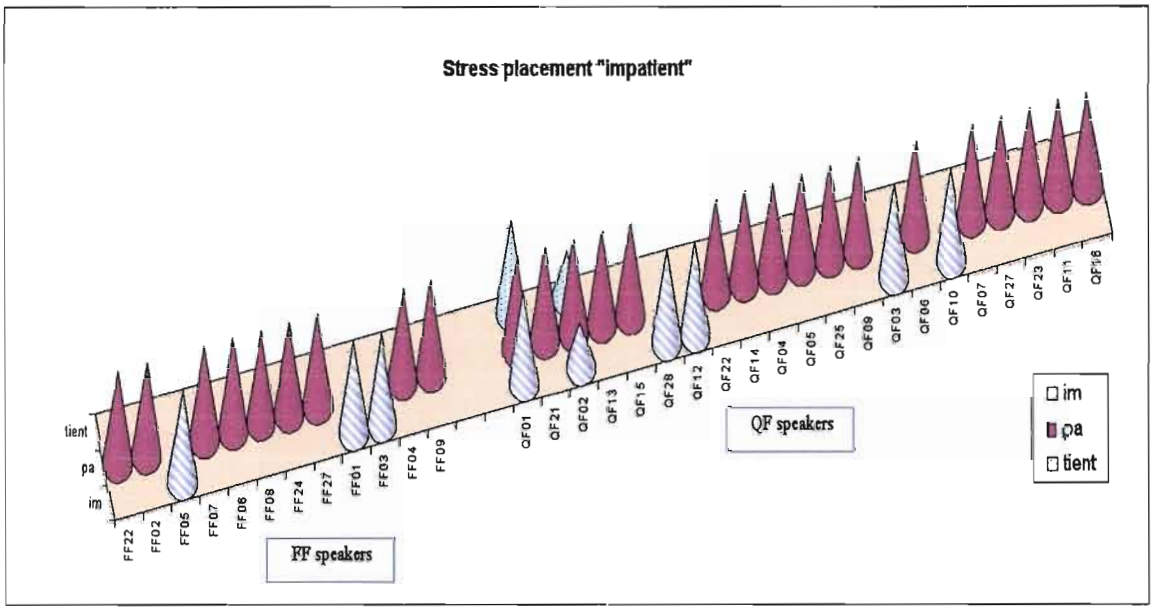


Figure 4.17: *Production of stress in "impatient" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

The word "motivated" was adequately produced by 4 of the 12 FF speakers. Of the remaining 8 FF speakers, 6 stressed the third syllable ("va") and 2 stressed both the first and the third syllable.

Among the 21 QF speakers, 4 stressed the first syllable, 13 stressed the third syllable ("va") and 4 stressed both the first and the third syllable. These findings are summarized in figure 4.18 below.

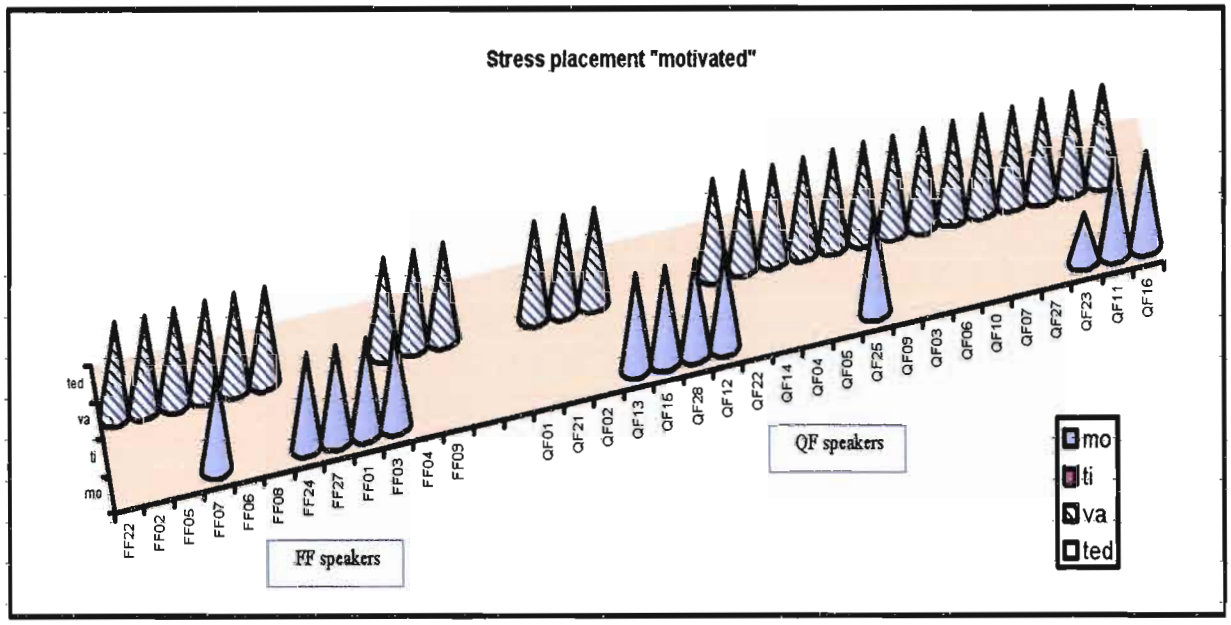


Figure 4.18: *Production of stress in "motivated" by FF and QF speakers. Plain colour cones represent syllables that, according to stress placement rules, should be stressed. Patterned cones represent syllables that should not be stressed; their presence therefore signifies a stress error. Large- scale figures available in Appendix 5.*

4.4.5 Discussion and conclusion

Lexical stress

The results obtained for the words "prison" and "offered" match predictions of incorrect stress placement either on the second syllable or on both syllables. The findings are however not consistent because only 3 out of 33 subjects stressed "prison" incorrectly while almost half stressed the second syllable – with or without the first – of "offered". Furthermore, the two subjects who stressed both syllables of "prison" equally are the lowest ranked of all francophone subjects, so it could be assumed that only very low competence subjects might stress this word incorrectly. The difference of success rate in the production between these two words might be explained by the nature of the words: "Prison" is a noun while "offered" a verb. It would look like subjects have assimilated that two-syllable nouns are usually

stressed on the first syllable. As the stress pattern for verbs is less reliable since it is stressed on the root – which can be on the first or on the second syllable –, subjects may just not have learnt the stress pattern of this word.

The words "perfect" and "study", also two-syllable words, were adequately produced by the majority of speakers. Although "study", like "offered", is a verb, the subjects did not have difficulty in stressing them correctly. It is possible that most subjects would have encountered the word "study" more often than "offered", which would also partially explain the difference in production.

The word "opportunity" was correctly stressed by the vast majority (94%) of subjects. As this word is longer than all the others, this may appear surprising. However, it is very likely that subjects have assimilated – consciously or not – stress placement of words ending in "-ity" as there are many such words. It has indeed been found that the degree of exposure has an impact on assimilation. "Findings suggest that age and suffix frequency both play a role in children's awareness of stress placement". (Jarmulowicz, 2002)

Similarly, the word "impatient" was generally well produced. As the suffixes /(j)ənt/ and /(j)ən/ are very common, it is very likely that many subjects have assimilated stress placement on these words.

"carpenter" however was not easily dealt with by these non-native subjects. This can also be explained by the particularities of English stress rules. Indeed, words ending with the suffix "er" are stressed like the word it is attached to (i.e. the suffix does not influence stress placement). For instance, "advise" gives "adviser", "condition" gives "conditioner". As the root of "carpenter" is not obvious, subjects who have not encountered the word enough cannot guess how to stress it.

Finally, the results for the word "motivated" show that the majority of subjects stressed it incorrectly. Moreover, the mistake was always the same, that is to say, speakers stressed the third syllable. This can be explained first, by the scarcity of words stressed like "motivate", on the antepenultimate syllable and second, by the influence of words in "-ation". As the latter, found in abundance in English, are stressed on the penultimate syllable (ex: motivation), non-native speakers produce the same stress pattern on the corresponding "-ate" word.

Sentence rhythm

The result of 85% of speakers having stressed "him" is quite considerable compared to the percentage of the errors for other words. However, the reason for this unwarranted stress is related to segmentation. The word "him" is stressed because it is the AP final syllable.

From the above results, it would seem that the five hypothesized stress placement did not all materialize in the data. Specifically, little evidence of final syllable stress – alone or combined with another stress – was uncovered. The next section however provides information that helps explain this situation.

Instrumental analysis

The results reported in the auditory-perceptual analysis part do not support the hypothesis that the realization of French stress (i.e. AP early and late rises) interferes with the production and placement of English stress.

The results of the instrumental analysis paint a different picture however. Indeed, many subjects produced these rises, especially the late AP rise.

Figure 4.19 shows that the speaker produced a rise throughout the word "opportunity". While the judges detected these rises, they did not perceive them as stresses. They considered them to be similar to the "uptalk" phenomenon (see

Prosody: functions and uses). The judges felt that, while these rises – like uptalk – were somewhat distracting and annoying, they did not interfere with the perception of stress.

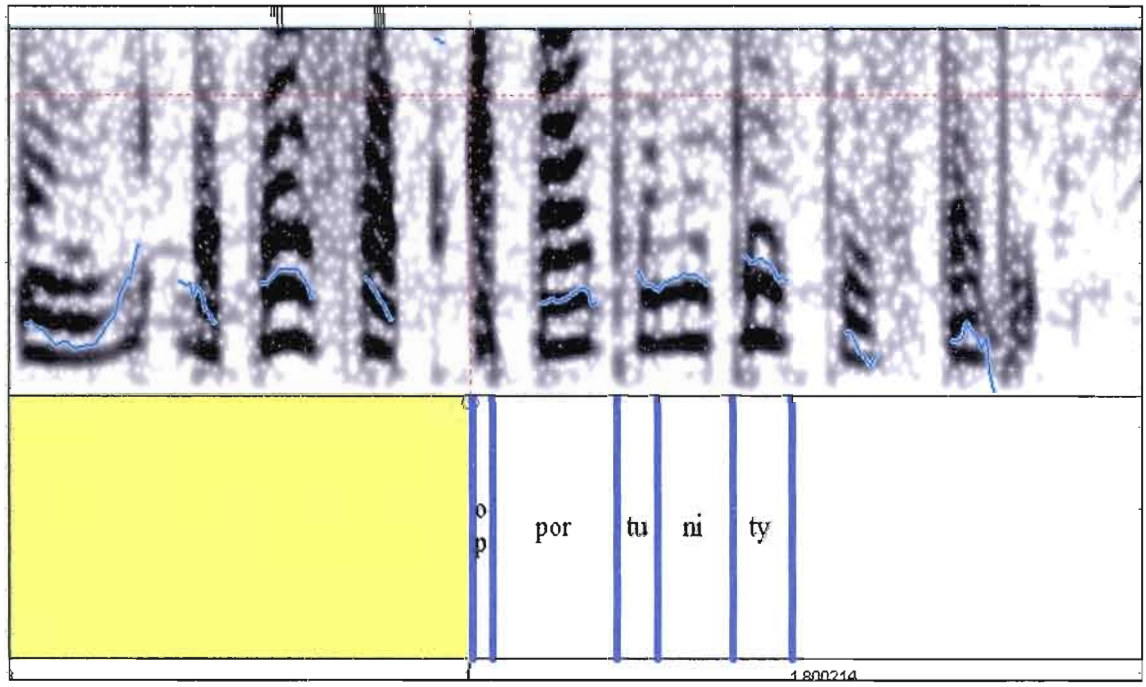


Figure 4.19: *Example of progressive F0 rise in the production of "opportunity" by a QF speaker.*

These rises materialized as high pitch, high intensity and considerable length. As these correlates are the recognized correlates of stress in English, it is surprising that the judges did not consider the syllables to be stressed. Some examples are provided below to illustrate and explain the perception of stress in different cases. To better appreciate these examples, it must however be understood that the francophone subjects often attributed the different acoustic correlates of stress to different syllables. For instance, greater frequency could be found on one syllable, greater intensity on another and greater duration on yet another. Contrary to native

speakers, it was rarely the case that all stress correlates were concentrated on the same syllable[41].

As can be seen from figure 4.20 below as well as from figure 4.19 above, in the production of "opportunity", the vowel sound of "tu" is shorter and of lower intensity than "ni" and "ty", and the maximum pitch is the same as that of "ni" and lower than "ty". The only justification for native speakers to identify "tu" as the stressed syllable of the word is that the contour is falling while that of the other two is rising.

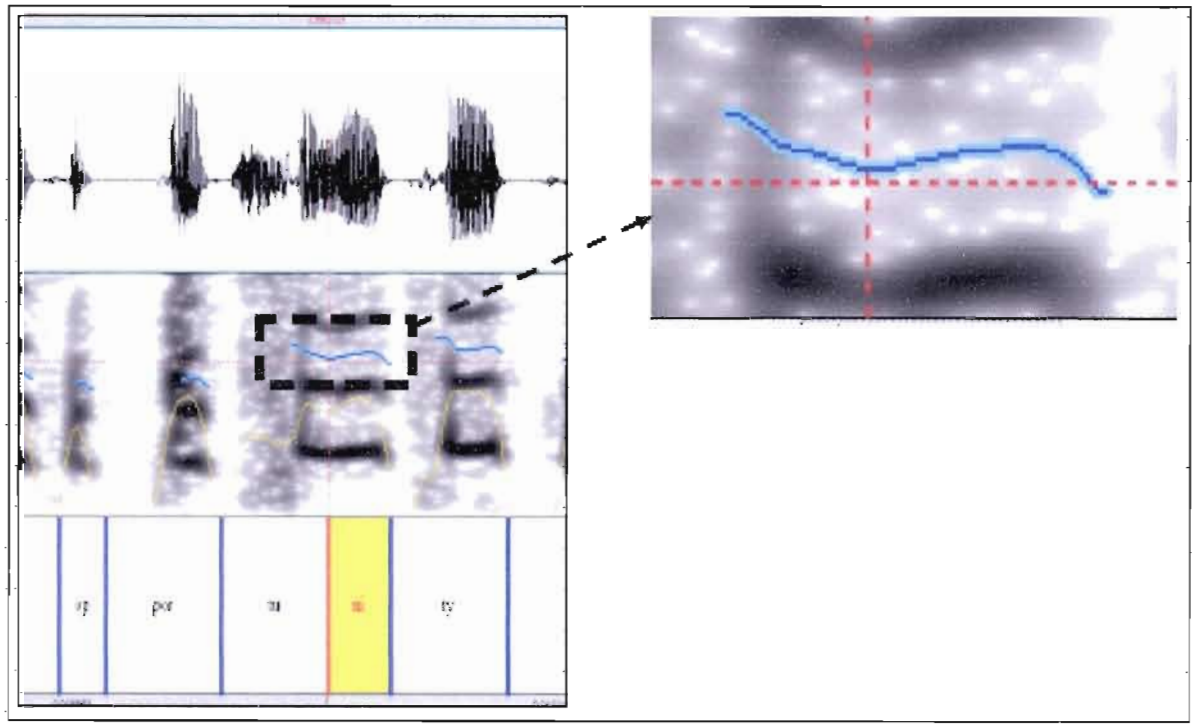


Figure 4.20: *Production of the word "opportunity" by a QF speaker. [Right] falling F0 contour over "tu".*

[41] Because of this, the auditory-perceptual analysis of these speakers' production was much more difficult than it would have been for the production by native speakers. This also explains why, where one judge would have likely sufficed for the analysis of native speakers' production, the number of judges to participate in this analysis was increased to three.

Yet, rising contours can be associated with stressed syllables. In figure 4.21 we can see that the speaker stressed the first two syllables of "carpenter", in such a way that it sounds like the nonce compound "car-penter".

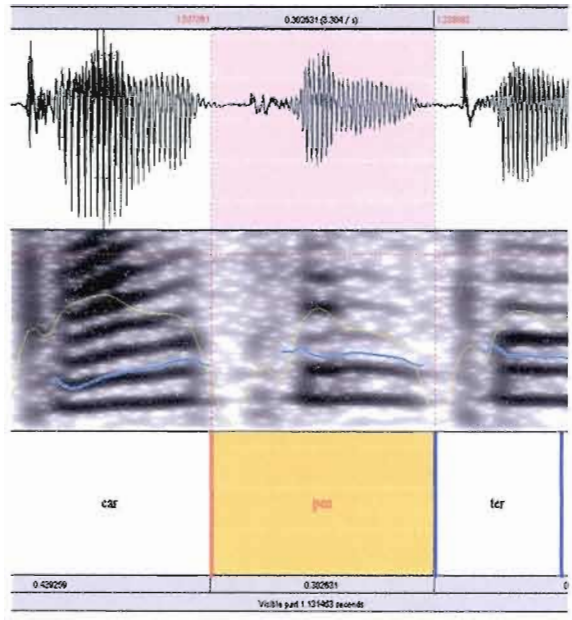


Figure 4.21: *Production of the word "carpenter" by a QF speaker.*

In fact, a same contour may be judged stressed or not depending on the situation. In figure 4.22 below, we can see that "son" and "him" are very similar in terms of intensity, frequency, contour and duration. Yet, the judges associated "him" with stress but not "son", probably because "pri" of "prison" was more stressed than "son" while "him" should just not be stressed at all.

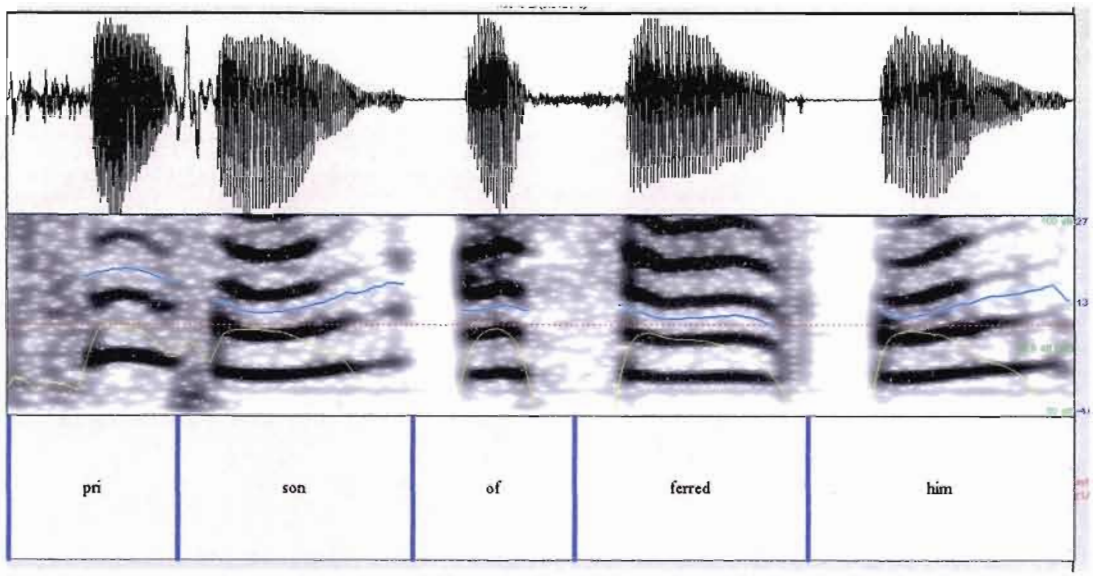


Figure 4.22: *Production of the segment "prison offered him" by an FF speaker.*

Finally, as can be seen from figure 4.23, sometimes, a greater amount of energy (indicated by darker lines on spectrogram) may be the feature that indicates that the syllable is stressed although the other correlates would indicate otherwise.

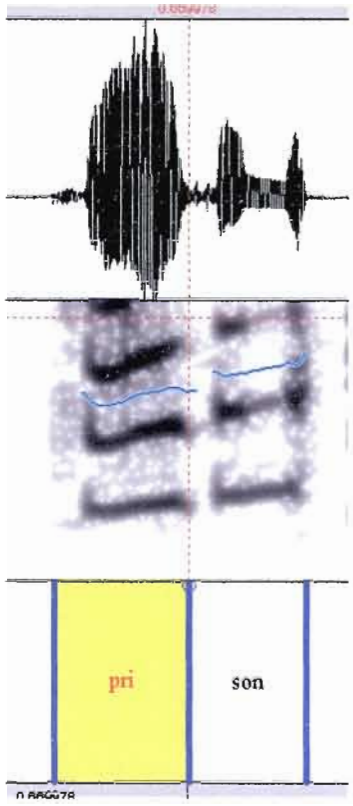


Figure 4.23: *Production of the word "prison" by a QF speaker.*

It must also be mentioned that the subjects attributed the different acoustic correlates of stress to different syllables. In this sense, greater frequency may be found on one syllable, greater intensity on another and greater duration on yet another. As specified in footnote 35, stress correlates being spread across different syllables made it difficult to identify stress placement.

We can therefore understand that the perception of stress is more complex than is generally assumed. Factors such as frequency, intensity, duration, pitch contour and concentration of energy are all to be taken into account. Establishing the matrix of relative importance of these correlates in the perception of stress is however beyond the scope of this work.

The results of this instrumental analysis do indicate that, although they were not recognized as stress, final AP rises were produced by native speakers of French. These rises were generally produced in greater numbers by lower level subjects. The phenomenon was amplified by the fact that they segment the sentences into more APs. Still, some higher level subjects, whose rhythm was otherwise quite suitable, persistently produced them.

Conclusion:

From the lexical analysis, we saw that our FF and QF subjects had assimilated the basic stress placement rules, as the vast majority adequately stressed the words "prison", "perfect", "opportunity", "study", "impatient". Knowledge of the more difficult rules (suffix "-ate") and of non-rule based stress (ex: "carpenter") did not increase with speaker ranking.

The relatively poor achievement of adequate stress on "offered" might indicate that a rule offering variety (the root of a verb can be on the first or the second syllable) and level of exposure to a word are relevant.

From the poor success rate of not stressing "him", we can suggest that phonological sentence rhythm is more difficult to acquire than lexical stress.

While the subjects managed to place stress on most words adequately, their phonetic production of stress and of the unstressed syllable was much less reliable. The acoustic correlates used to describe stress and used by native speakers in the production of the stressed syllable often were produced on different syllables by the francophone subjects.

This experiment also showed that the assumptions that some of the speakers would not place stress correctly and that some might produce a final rise were confirmed.

For instance, the word "carpenter" alone was produced in enough ways to confirm four of the five hypotheses. Indeed, 4FF and 2QF stressed the second syllable which confirms hypothesis 3; of these speakers, some produced a final rise confirming hypothesis 2. Also 2FF and 10QF correctly stress the word, thus confirming hypothesis 5; of these speakers who stressed the word correctly, some produced a final rise hypothesis 4. Hypothesis 1 was not confirmed on long words but was on words of two syllables.

The assumed impact of a final rise on perception of stress by native speakers was however incorrect. Indeed, it was determined that native speakers of NAE do not associate final rises with stress. This does not mean that rises are not a mark of stress in French. It only means that a final rise is perceived as just that a final rise, not as stress.

The propositions made in 4.3.1.4.3. must therefore be corrected as follows:

1. ~~Stress will not be produced on the normally stressed syllable but stress will be placed on the last syllable.~~
2. Stress will be placed on the wrong syllable ~~as well as on the last syllable~~ and a final rise will be produced.
3. Stress will be placed on the wrong syllable ~~but not on the last syllable~~ and no final rise will be produced.
4. Stress will be placed on a correct syllable ~~as well as on the last syllable~~ and a final rise will be produced).
5. Stress will be placed on the correct syllable and no final rise will be produced.

This experiment also shows that perception of stress by native speakers is more complex than is usually assumed. While the perception of higher pitch, higher intensity, and greater duration certainly equates to the perception of stress as

produced by native speakers, it was shown here that contour and energy must also be taken into account. An understanding of how native speakers compute these correlates when they are distributed over a number of syllables requires further research.

4.5 Study 2: Standard Chinese

The purpose of this study is to analyze the production of English by native speakers of Standard Chinese. The areas to be analyzed are some of the potential difficulties stemming from the comparison of the grammars of NAE and of SC. The specific objectives of each experiment will be explained in the appropriate section.

4.5.1 Purpose of study

The purpose of this study is to analyze the production of English by native speakers of Standard Chinese. The areas to be analyzed are some of the potential difficulties brought forward from the comparison of the grammars of NAE and of Chinese. The specific objectives of each experiment will be explained in the appropriate section.

4.5.2 Subjects

Five native speakers of NAE, all teachers of ESL, provided the control recordings. The details about the speakers of SC are provided in the tables below[42].

Table 4.5 Native speakers of Standard Chinese

Number of subjects:	24
Number of male subjects:	5
Number of female subjects:	19
Average age:	36,3
Youngest subject:	25
Oldest subject:	44
Average years of ESL learning:	9

[42] More details about individual speakers of all groups are provided in Appendix 6.

4.5.3 Experiment 3

4.5.3.1 Objectives

Chen (2007) reports that "Chinese speakers learning Spanish seem to interpret Spanish stress as a combination of a rising and a falling tone". If speakers of a tonal language perceive tones in non-tonal languages, it may be possible that they also use tones in their production of non-tonal languages. In the case at hand, the question is "do native speakers of SC use SC tones in their production of English?" Certainly, phonologically speaking, pitch accents of English *are* tones. However, besides the fact that tones are lexical in SC but not in English, pitch accent and tones differ in their phonetic production. First, the phonetic pitch range of SC tones is greater than that of English rhythm.

> "[Chen's (1974)] results show that the average pitch range of the native Chinese speakers speaking Chinese was 1.5 times wider than that of the English speakers when they spoke English" (Wang et al, 2006)

Second, SC tones require a specific use of the tonal space while English pitch accents are adjusted by each speaker. The difference between speakers who use a narrow pitch range and those who use a wider pitch range in English is generally understood to reflect personality and mood (for instance, timid and bored for the former, outgoing and excited for the latter). In English, greater or narrower pitch range does not impede understanding though. In SC, not respecting the traditional use of tonal space creates problems. For example, learners of SC as a second language make as many tonal space errors as contour errors. Miracle (1989) found that L2 learners of SC produced tones 1 and 4 too low in the tonal space and tones 2 and 3 too high in the tonal space. This shows that in SC, tone pitch height and pitch range are standardized.

The objective of this experiment is thus to examine if native speakers of SC reproduce SC tones in their production of English. If so, a rising pitch accent would for instance be produced as a tone 2 and a falling one as a tone 4.

4.5.3.2 Procedure and Stimuli

As was explained in chapter 3 (see 3.4.5.1.2.), Chao (1930) described SC tones in their tonal space which he divided in five levels. This is convenient to compare the four tones but, as there is no reference to measurable correlates (fundamental frequency or semitones or other), relating tones to pitch accents is particularly challenging. For this reason, in this experiment, the speech of the same native speakers of SC will be compared for their production of pitch accents in English and their comparable tones in SC. This will be done by comparing the pronunciation of the English word "fan" in different contexts and the production of [fan] in SC.

The production of *fan* in English by native speakers of SC is analyzed in the four instances (underlined) in the sentence below:

There are three fans: a black fan, a white fan, and a red fan.

The production of [fan] in SC by native speakers of SC is analyzed in the twenty instances (underlined) in the sentences below[43]. These include five sentences with each word, that is to say the segment [fan] plus a tone (i.e. fān, fán, fǎn, fàn). To allow for the possibility of the phonetic production of a same tone to be different from its phonological representation according to its position in a sentence, each word was inserted in five sentences. In the first sentence (A), the word was sentence initial; in the second (B), mid sentence; in the third (C), affirmative final; in the fourth (D), pre-pause; and in the fifth (E), interrogative final.

Tone 1

A. Fān chuán yùn dòng hěn yǒu qù.
B. Wǒ xǐ huān fān chuán yùn dòng.
C. Yí dìng yào chēng zhù fān.
D. Bù dān yào chēng zhù fān, hái yào bǎ hǎo duò.
E. Nǐ néng bù néng chēng fān?

[43] The subjects were given the sentences in Chinese characters only. These are provided in Appendix 2. The pinyin version is used here for the benefit of non-tonal language readers.

Tone 2

A. Fán suǒ de shǒu xù ràng tā tóu téng.

B. Dēng jì shǒu xù de fán suǒ ràng tā tóu téng.

C. Gōngsī de shì ràng tā xīn fán.

D. Bù néng rěn shòu píng fán, jiù nán yǐ chéng gōng.

E. Tā shì bú shì hěn píng fán?

Tone 3

A. Fǎn chéng de piào yě dìng hǎo le.

B. Tā méi yǒu yù dìng fǎn chéng piào.

C. Zhè gè jià gé bāo kuò wǎng fǎn.

D. Rǔ guǒ bāo kuò wǎng fǎn, jiù bú suàn guì.

E. Tā mén shì bú shì yào móu fǎn?

Tone 4

A. Fàn chī wán le, cài hái yǒu yì xiē.

B. Tā hē wán jiǔ yǐ hòu, méi chī fàn jiù zǒu le.

C. Tā guāng chī cài, bù chī fàn.

D. Wǒ men qǔ chī fàn, hǎo bù hǎo.

E. Nǐ zuó tiān wǎn shàng chī de shén me fàn?

For each production of *fan* in English and in SC, pitch level was measured at the 10%, 30%, 50%, 70%, and 90% marks of the voiced part. The analysis was conducted with a program written to work with Praat. When the pitch level is outside the defined bracket or when there is no voicing - which happens with creaky voice -, the measure returns as "undefined". For the experiment, one of the measures returned as "undefined", the pitch was measured manually moving away from the target to the nearest reading. For instance, if the 30% mark returned as undefined, a measure was searched for at 29%, 31%, 28%, 32%, and so on. If no measure was available half way to the next target (on our example, 20% and 40%), no measure was entered.

4.5.3.3 Hypothesis and Analysis

Hypothesis: SC subjects will produce the tone whose contour most closely matches the contour of the intended English pitch accent.

More specifically, SC speakers will transfer tones 2 and 4 to produce, respectively, rising and falling pitch accents. This is based on the same principle as Flege's (1995) Speech Learning Model which specifies that similar L1 and L2 features are not as easily acquired as features that are very different. According to the same principle, tone 1 and tone 3 (in the LH form) are unlikely to be transferred into English because they are very different from any English pitch accent.

Looking more specifically at the sentence at hand, we can see (as illustrated in figure 4.24) that native speakers produce:

- the first instance of "fan" (henceforth "fan1") as a falling tone
- the second instance of "fan" (henceforth "fan2") as a high rising tone
- the third instance of "fan" (henceforth "fan3") as a high rising tone
- the fourth instance of "fan" (henceforth "fan4") as a low falling-rising tone

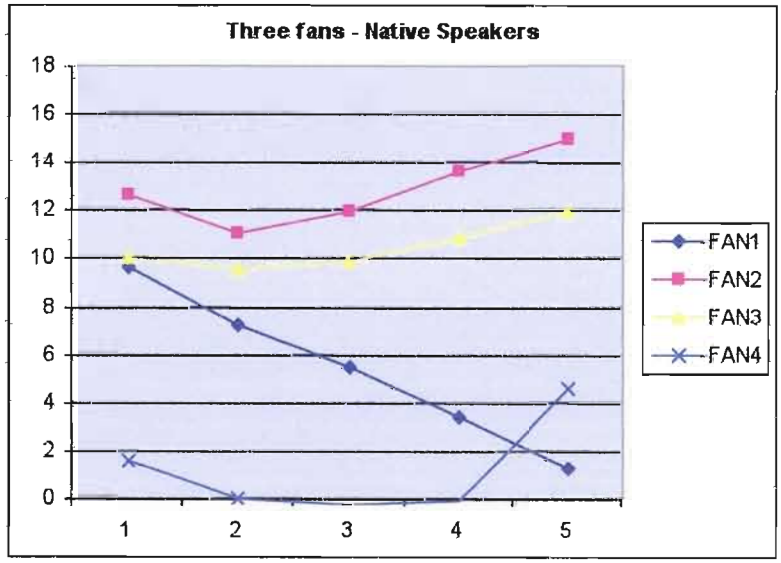


Figure 4.24: *Experiment's SC native speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the syllable "fan" in the four occurrences of the word in the sentence.*

Bearing in mind the shape of SC tones (reproduced below) and the contours produced by native speakers (figure 4.24 above), one could therefore hypothesize that native speakers of Standard Chinese would produce:

Fan1 as a 4th tone,

Fan2 as a 2nd tone

Fan3 as a 2nd tone

Fan4 as a 3rd tone

	AM notation	Tone shape
1st tone	H	High
2nd tone	LH	High rising
3rd tone	L(H)	Low falling (rising)
4th tone	HL	High falling

Analysis:

The combined SC speakers produced the following contours for the four words:

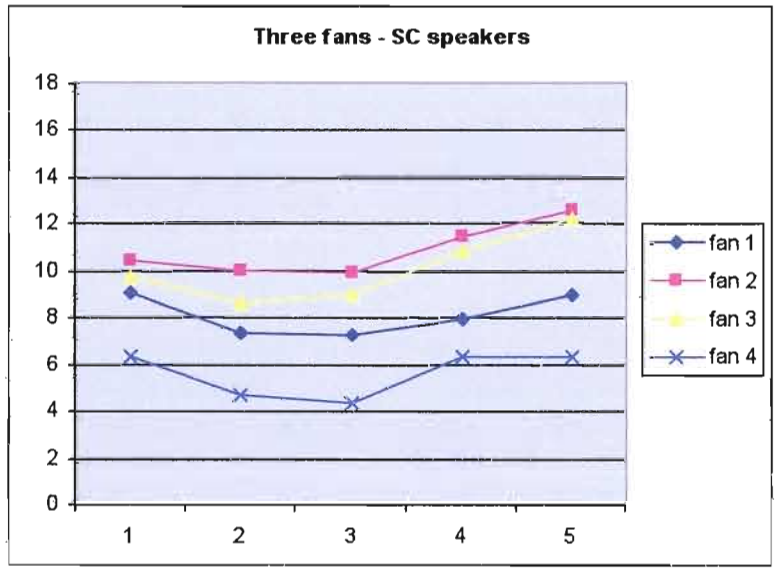


Figure 4.25: *SC speakers average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the syllable "fan" in the four occurrences of the word in the sentence.*

Comparing the production of native speakers of English and of SC, we can see that:

- fan1 was produced as a falling tone by NAE speakers and by 4 SC speakers (ranked 11th, 14th, 15th and 23rd) but as a rising tone by the other 20 SC speakers.
- fan2 was produced as a rising tone by both groups
- fan3 was produced as a rising tone by both groups
- fan3 was produced a little lower than fan2 by both groups
- fan4 was produced as a low falling-rising tone by both groups
- fan4 was produced lower than the other three occurrences of "fan" by both groups
- fan4 was produced considerably lower by NAE speakers than by SC speakers

The first observation is that, for fan1, the great majority (83%) of native speakers of SC did not use a falling tone but a rising tone. This shows that SC speakers do not generally associate a falling contour with a phrase final accent. SC speakers' production of fan1 will be compared to a rising SC tone.

Figures 4.26 to 4.29 show the production of fān (tone 1), fán, fǎn, fàn respectively. Each graphic includes the combined production of all SC speakers for the tone in the different situations: A: sentence initial; B: mid sentence; C: affirmative final; D: pre-pause; E: interrogative final.

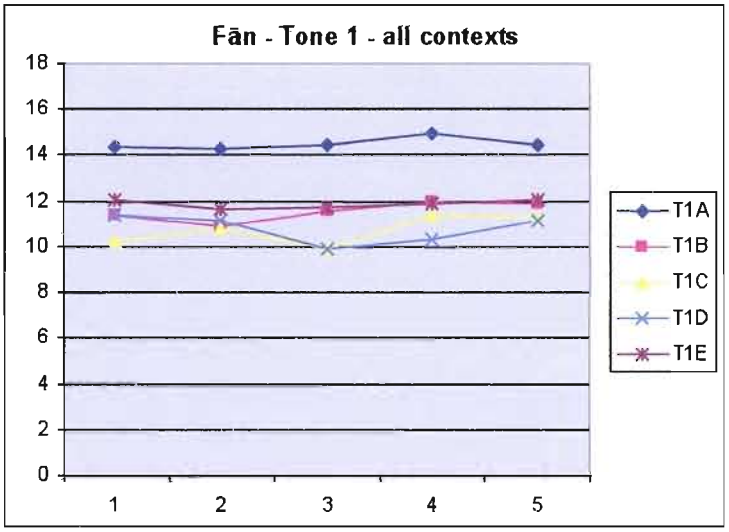


Figure 4.26: *SC speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the word "fān" in five sentential positions.*

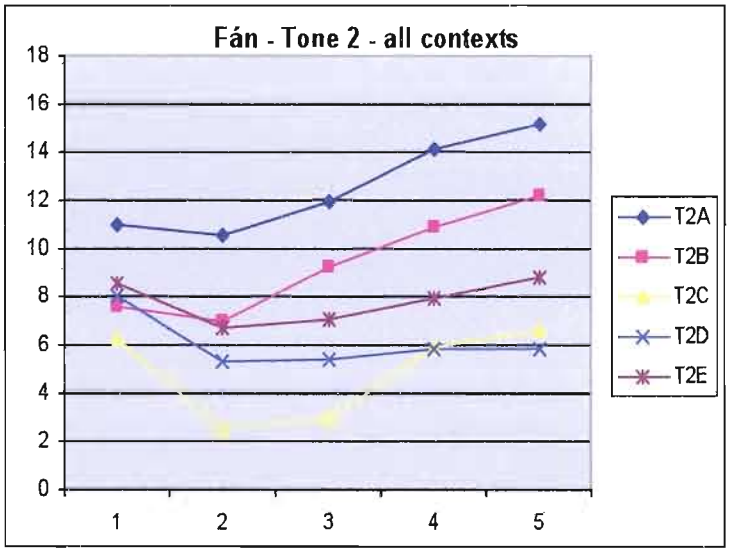


Figure 4.27: *SC speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the word "fán" in five sentential positions.*

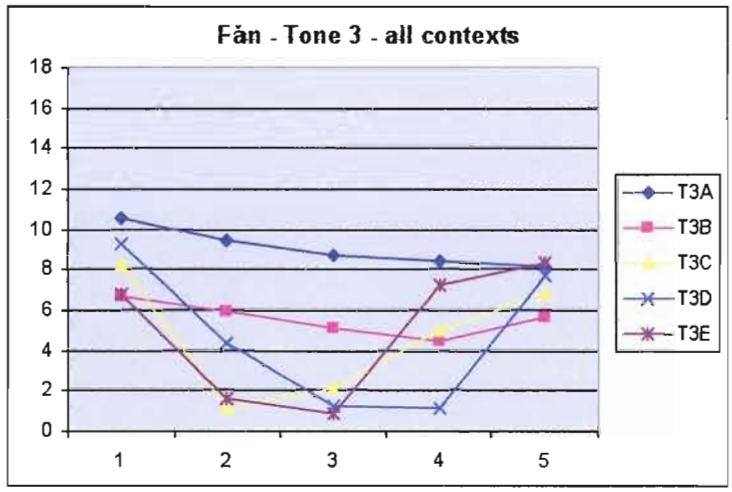


Figure 4.28: *SC speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the word "fǎn" in five sentential positions.*

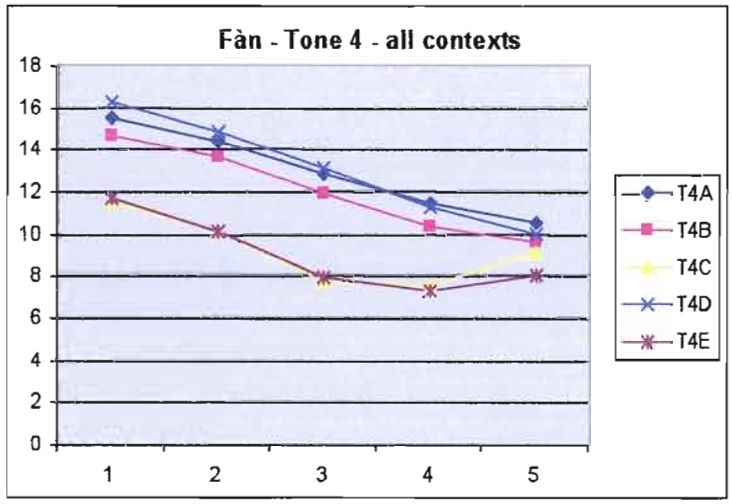


Figure 4.29: *SC speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the word "fàn" in five sentential positions.*

Although pitch level varies according to the place of the segment in the sentence, the shape of the tone varies very little. The exception is tone 3 which is a falling tone in situation A and, with a slight rise at the end, in situation B but a clearly falling rising

tone in all pre-boundary positions (i.e. C, D, and E). It must also be noted that the falling rising production of tone 3 is particularly difficult to measure because of the creakiness that accompanies the lowest frequencies.

As all instances of "fan" in the English sentence produced by SC speakers were produced as a rising contour, they will be compared with tone 2, the SC rising tone. The different productions of "fan" and of tone 2 are presented in the same graphic (figure 4.30).

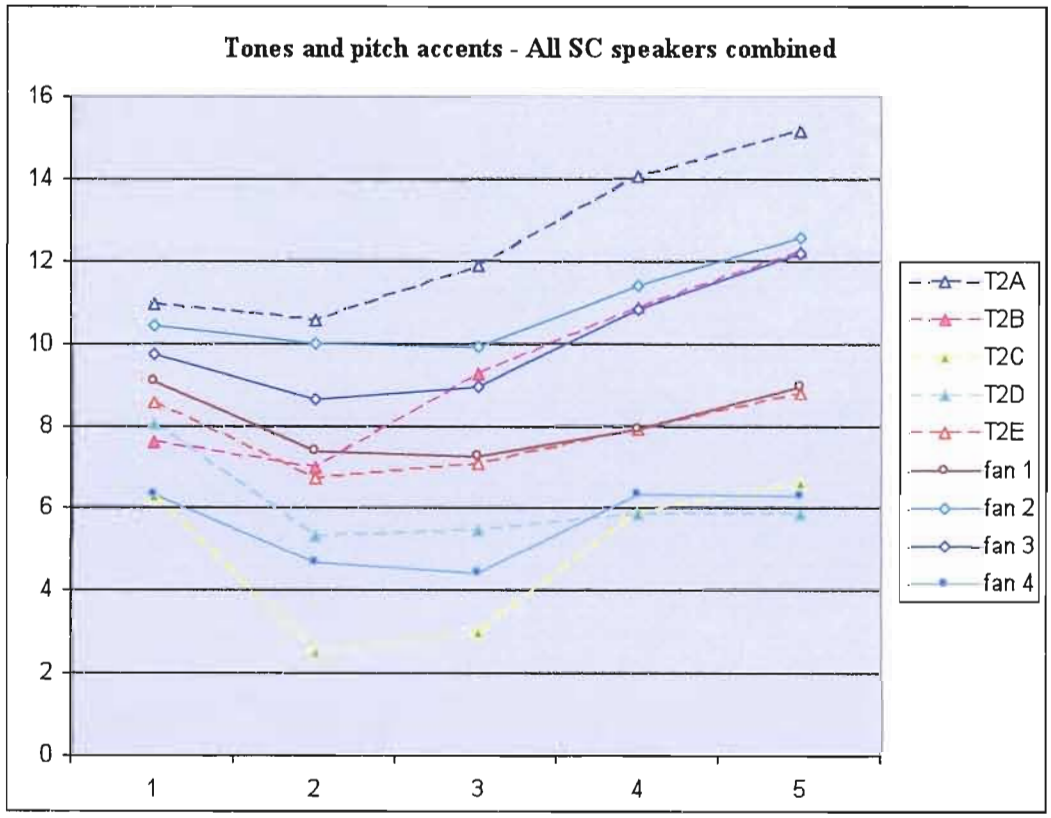


Figure 4.30: *SC speakers' average F0 in semitones at 10%, 30%, 50%, 70%, 90% of the Chinese word "fán" in five sentential positions and of the four occurrences of the English word "fan" in the sentence "There are three fans: a black fan, a white fan, and a red fan.*

We can see that contours are quite similar. More specifically, the contour of fan1 is almost identical to that of T2E, the production of tone 2 on an interrogative final syllable.

Fan2 and fan3 have the same contour as, but are higher than, fan1.

Fan4 has a contour similar to T2C (affirmative final).

The use of the tonal space is also more similar between tone 2 and the production of "fan", than between the other tones and "fan", as illustrated in table 4.2.

Table 4.6: Tonal space used for each tone and for the English word "fan".

(in semitones)	Fan		Tone 1		Tone 2		Tone 3		Tone 4	
	beginning	end	beginning	end	beginning	end	beginning	end	beginning	end
Highest	10,43	12,6	14,4	14,49	10,96	15,1	10,51	8,34	16,2	10,5
Lowest	6,34	6,29	10,23	11,17	6,27	5,83	6,73	5,64	11,71	8
Range	4,09	6,27	4,17	3,32	4,69	9,3	3,78	2,7	4,49	2,54

The conclusion is thus that, as a group, SC speakers produced the English pitch accents as varieties of tone 2. One is entitled to wonder though if these results are just the effect of combining all speakers' production.

Confirmation of the relation between tones and pitch accent contours in SC speakers' production is made clear from the speech idiosyncrasies of some of the speakers.

In figures 4.31 and 4.32, we can see that speaker SC12 has a tendency to produce a flat or slightly falling tone2 in position D. This is also observed in SC12's production of fan4.

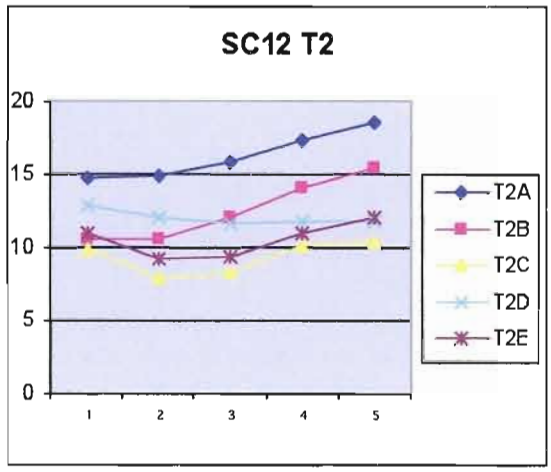


Figure 4.31: *SC12 speaker idiosyncrasy in the production of pre-pause tone 2.*

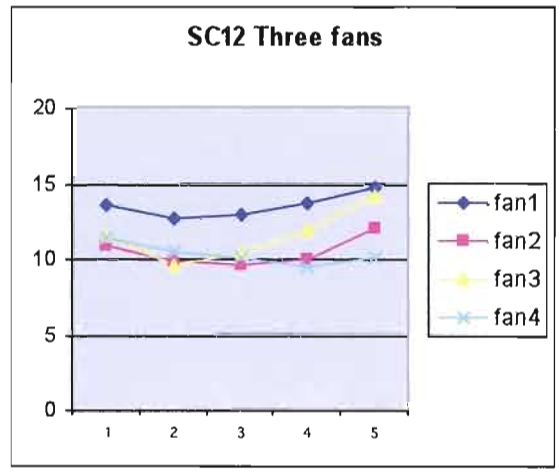


Figure 4.32: *SC12 speaker idiosyncrasy in the production of the English word "fan" in position 3.*

Similarly, we can see from figures 4.33 and 4.34 that speaker SC17 produces Tone2 very low in the F0 range and does the same thing for fan1 and fan3. This shows that the speaker produced the first phrase ("*there are three fans:*") as a complete sentence as T2C and T2E are sentence final syllables. The same speaker also produces tone 2D with a falling F0 contour which she also does in fan4. This production of tone 2 must not be mistaken for a tone 4, which the speaker produces quite differently (figure 4.35).

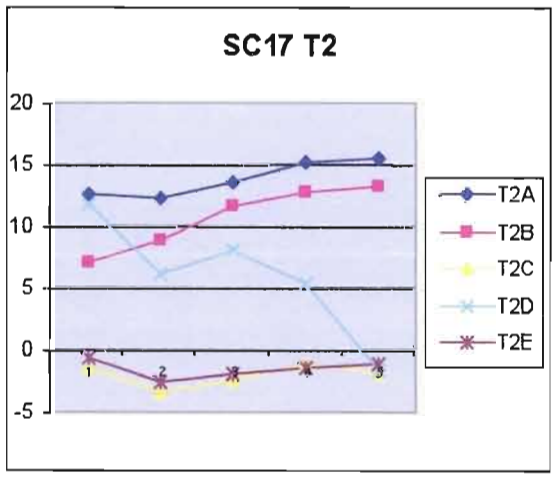


Figure 4.33: *SC17 speaker idiosyncrasy in the production of pre-pause tone 2.*

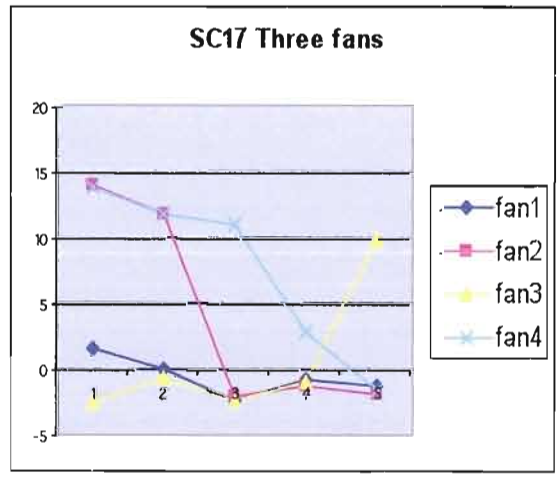


Figure 4.34: *SC17 speaker idiosyncrasy in the production of the English word "fan" in position 4.*

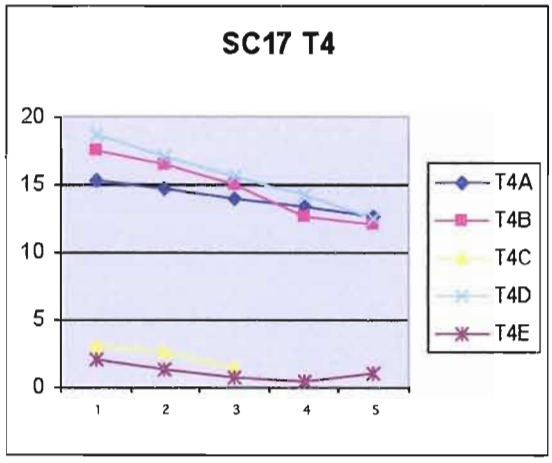


Figure4.35: *SC12 speaker's production of Chinese "fān".*

The majority of subjects produced tone 2 (in all positions) with a short fall (between the 10% and 30% marks) before a rise. This tendency to first produce a contour contrary to the intended direction is specific to tone 2. This is also visible in the production of the four "fan" pitch accents.

Speaker SC10 data (figures 4.35 and 4.36) illustrate this tendency. This same speaker has the idiosyncrasy of keeping a smooth contour throughout the tone up to the 70% mark before altering direction. It gives the impression of a break at the 70% mark. For instance the last segment of T2D is flat, while those of T2A and T2B show a slower increase. This same idiosyncrasy is also obvious in the subject's production of fan1, fan2, and fan3.

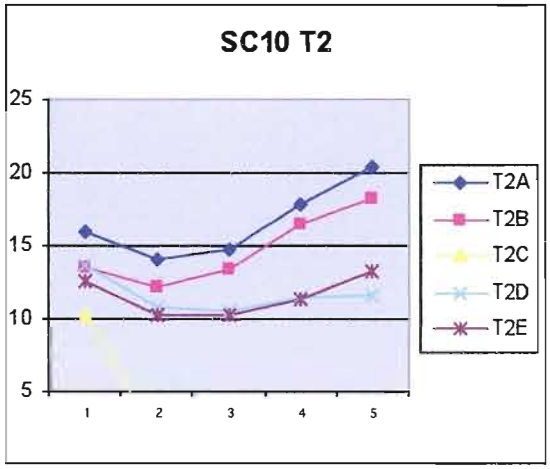


Figure4.35: *SC10 speaker idiosyncrasy in the production of the last section of tone 2.*

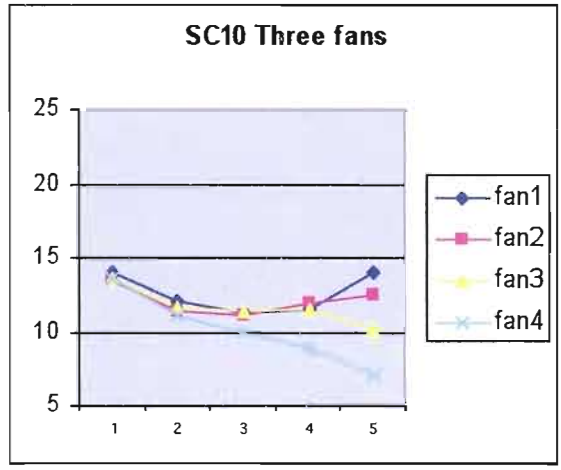


Figure4.36: *SC10 speaker idiosyncrasy in the production of the last section of the English word "fan" in positions 1, 2 and 3.*

One subject did not use Tone 2 – or tone 2 like – contours to produce the four pitch accents. In figures 4.37 and 4.38, fan2, and fan3 are produced as tone1E by the speaker.

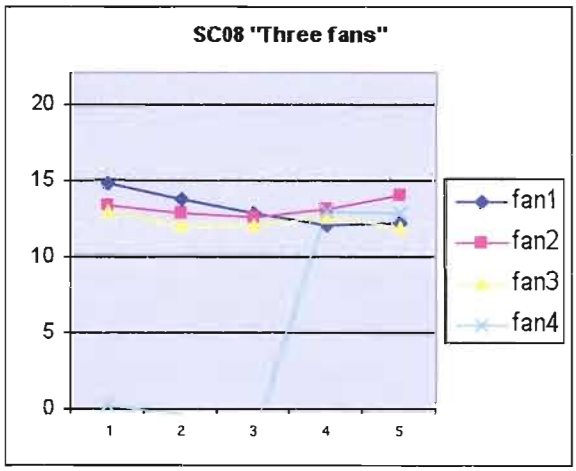


Figure 4.37: *SC08 speaker production of the English word "fan" in all positions.*

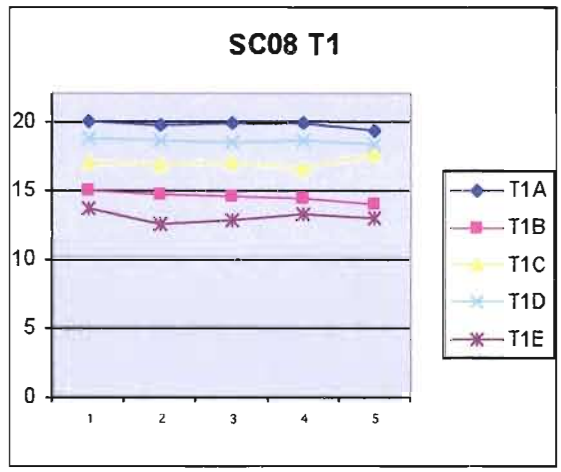


Figure4.38: *SC08 speaker production of Chinese tone 1.*

Five SC subjects (see 4.36 above and figures 4.39, 4.40, 4.41 and 4.42 below) used a falling pitch accent in the production of fan4, like the native speakers did. The last three of these subjects also produced fan1 as a falling contour like native speakers of English did.

In view of the very limited number of subjects who produced a sentence final falling pitch accent, it could be said that sentential prosody is acquired quite late but before list introductory phrase boundary tone.

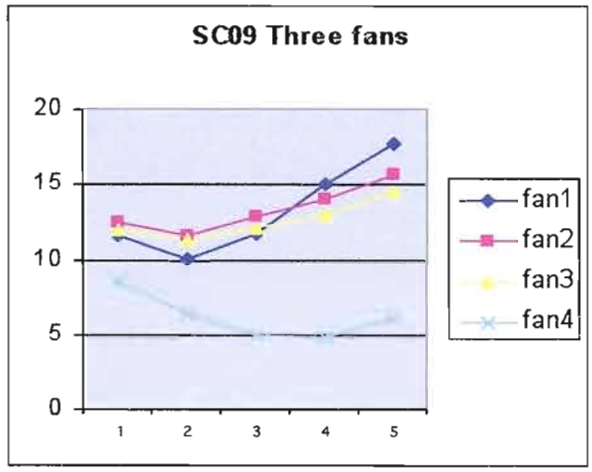


Figure 4.39: *SC09 speaker production of a falling pitch accent for* fan4.

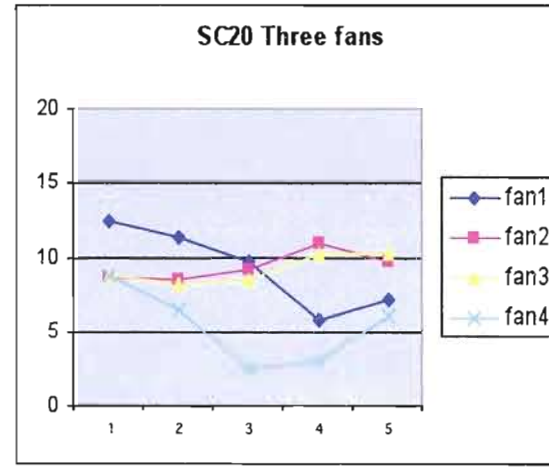


Figure 4.40: *SC20 speaker production of a falling pitch accent for* fan4 *and* fan1.

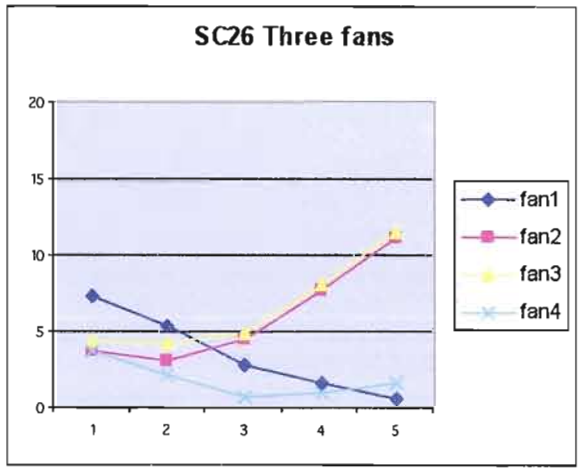


Figure 1.41: *SC26 speaker production of a falling pitch accent for* fan4 *and* fan1.

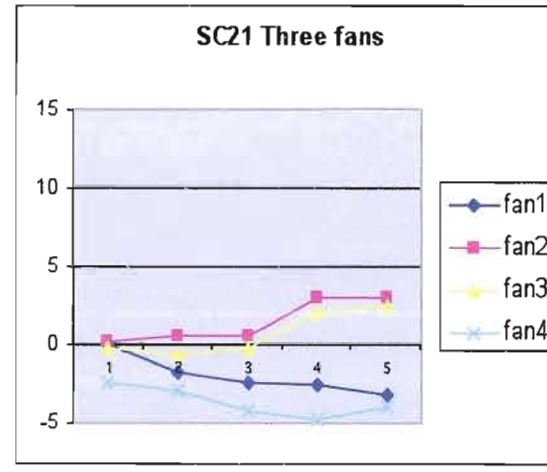


Figure 4.42: *SC20 speaker production of a falling pitch accent for* fan4 *and* fan1.

The falling tone used by these speakers is similar to the production of a SC tone 3 (low falling) rather than that of a tone 4 (high falling) as can be seen from figures 4.43 to 4.45.

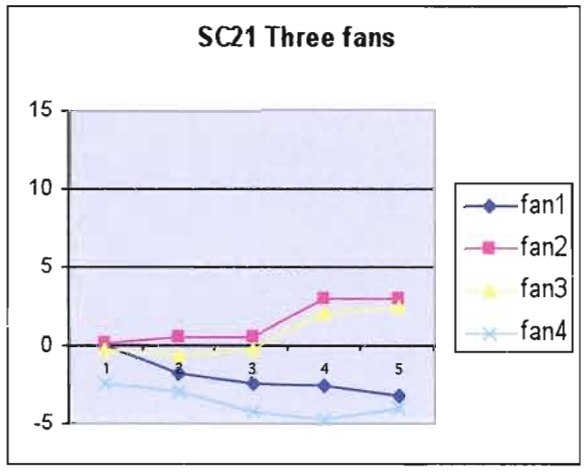


Figure 4.43: *SC20 speaker production of a falling pitch accent for* fan4.

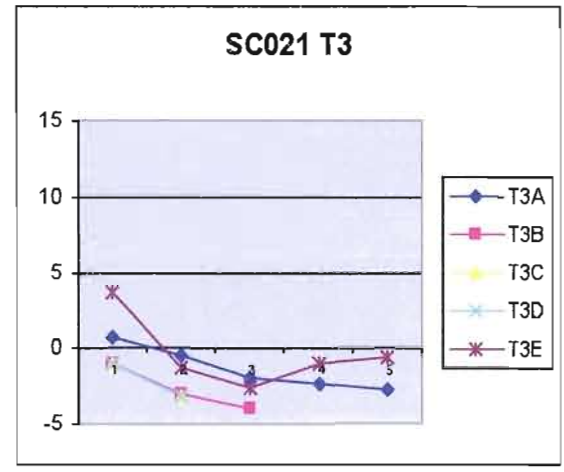


Figure 4.44: *SC20 speaker production of a tone 3 in all sentential positions.*

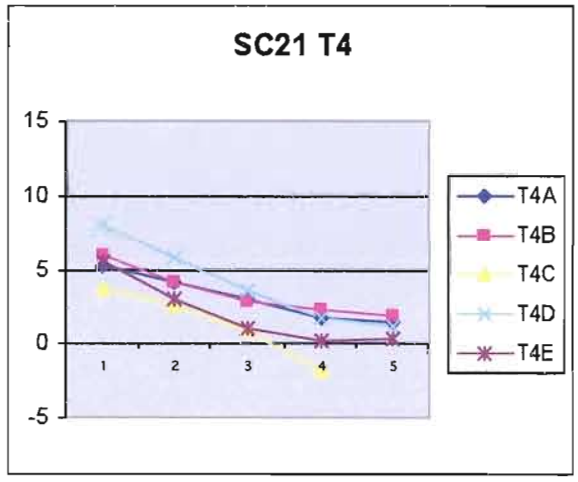


Figure 4.45: *SC20 speaker production of a tone 4 in all sentential positions.*

No clear relation between the production of falling tones for fan1 or fan4 and the subjects ranking was found as they were ranked 11th (SC20), 13th (SC08) 15th (SC21) 22nd (SC09) and 23rd (SC26) out of 24 speakers.

4.5.3.4 Conclusion

The hypothesis was that SC subjects would produce English pitch accents as the SC tones whose contour matched most closely the contour of the intended pitch accents.

The experiment confirmed this hypothesis insofar as SC speakers produced the contour and pitch range of tone 2 to produce all rising pitch accents.
From the SC subjects' production of the segment [fan] plus a tone (i.e. fān, fán, fǎn, fàn), it was seen that the pitch level of a same tone varies according to the place of the segment in the sentence. From this, it was further observed that certain productions of the Standard Chinese segments were more closely associated with certain productions of English "fan". For instance, the contour of fan1 is closely associated with T2E (interrogative final) and that of fan4 is related to T2C (affirmative final).
Fan2 and fan3 have the same contour as, but are higher than, fan1.

The connection between tone 2 and phrase final pitch accents was confirmed by the fact that SC speakers reproduce their tone 2 idiosyncrasies in their production of "fan".

During this experiment, it was also observed that most SC produced all instances of "fan" as high rising contours while NAE used a high falling contour for fan1. This discrepancy suggests that most SC speakers have not assimilated the boundary prosodic contour produced at the end of a phrase that introduces a list. The few speakers who did produce a falling contour on fan1 or fan4 seem to produce a low falling tone 3 rather than a high falling tone 4. There were however too few examples for this research to be conclusive in this regard.

4.5.4 Experiment 4

4.5.4.1 Objectives

The objective of this experiment is to compare how native speakers of English and native speakers of SC align pitch accents with syllables.

4.5.4.2 Procedure and Stimuli

Procedure: Chinese sentences produced by SC speakers and English sentences produced by native speakers of English and by SC speakers were investigated using instrumental techniques. The sentences were segmented with the Praat software so as to indicate all syllable boundaries. Perceptual and visual assessment of contours as well as tones and pitch accent alignment leads to the analysis presented below.

Stimuli:

The English sentence analyzed is:

The old man holds a fan in his right hand.

The following SC sentences are analyzed:

fàn chī wán le, cài hái yǒu yì xiē

yí dìng yào chēng zhù fān

4.5.4.3 Hypothesis and Analysis

In English, pitch accents are phonologically associated with stressed syllables (figure 4.46a) but are not necessarily phonetically aligned with them (figure 4.46b).

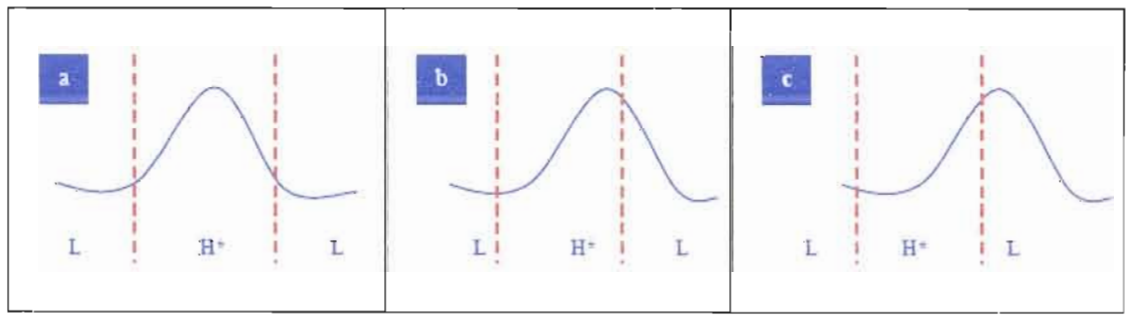


Figure 4.46: *Stressed syllable (a) phonetically aligned with phonological pitch accent, (b) phonetically delayed within stressed syllable boundaries, (c) delayed into next syllable.*

One of the manifestations of this phonetic non-alignment is F0 peak delay, that is to say, the production of a stressed syllable's F0 peak on the next syllable (figure 4.46c.) This phenomenon has been observed, among others, in English (Silverman and Pierrehumbert, 1990), in Greek (Arvaniti and Garding, 2007) and in Mexican Spanish (Prieto, van Santen, and Hirschberg, 1995).

It is generally assumed that the syllable after the stressed syllable needs to be unstressed to accommodate the delayed peak. This is confirmed by Chiang and Chiang (2005) who define the phenomenon as "the postponement in various prosodic conditions of an F0 peak from an accented syllable into the unaccented syllable directly following it."

As SC assigns a tone to the vast majority of syllables, and that the change of tone on a syllable amounts to a change of lexical item, it would seem unlikely that in SC an F0 peak could be delayed onto the next syllable. However, the absence of stress on the next syllable is not necessary. Indeed, Xu (2001) showed that tone 2 peak was regularly delayed at all speech rates and that tone 4 was delayed at fast speech rate. His experiment design placed these tones between two 3rd tones (without a final rise). The delayed peak was therefore produced at the beginning of an L tone.

Xu specifies that tone 2 peak delay, the most common, is mainly due to articulatory factors:

> Articulatorily, if a sharp F0 rise is implemented in the final portion of a syllable in the R[ising] tone, and when the following tone is L, a sharp turn needs to be made at the syllable boundary. To produce such a sharp turn, the larynx needs to first stop the pitch raising gesture and then start a pitch lowering gesture. This process should take time, and the result would be that the peak (i.e., the turning point) actually occurs in the following (L-carrying) syllable. (Xu 2001:29)

Lack of time to produce an abrupt change of gesture direction also explains tone 4 peak delays as it occurs mainly in fast speech.

Figure 4.47 below shows the syllable/tone alignment in the production of an SC sentence by one of the SC speaker taking part in the experiments. The arrows have been superimposed to show the theoretical direction of the tone. We can see that the phonetic realization of the sentence is very faithful to its phonological organization as each tone is confined within syllable boundaries. Figure 4.48 illustrates Xu's finding of a tone 2 delay before a tone 3 syllable. In fact, the delay applies to the tones that follow tone 2 rather than tone 2 as such. Indeed, we can see that the rising part of the rising tone of "hái" is accurately completed before the next syllable. However, as specified by Xu, the necessary gesture to bring F0 back to tone 3 level (very low) impedes on the production of "yǒu". In fact, we can see that this delay has a knock-on effect, that is to say, that all subsequent tones are shifted to the right. This is illustrated by the dashed oblique lines.

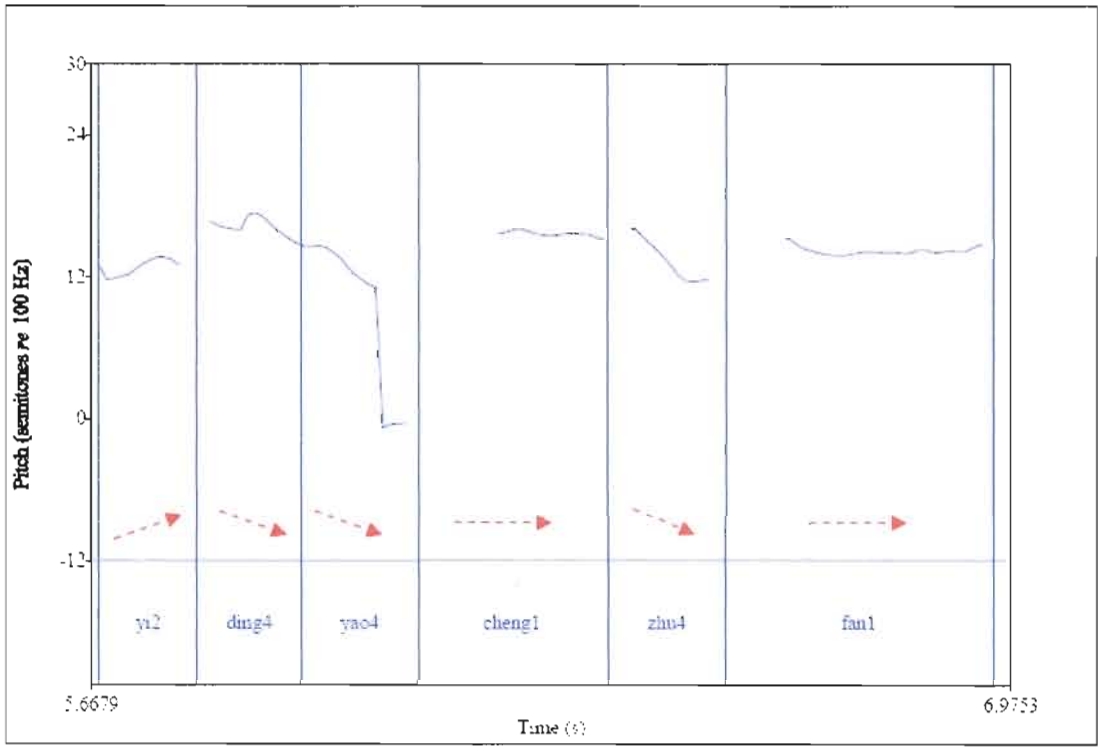


Figure 4.47 *Syllable/tone alignment in the production of an SC sentence by one of the SC speakers taking part in the experiments. The arrows show the theoretical direction of the tone. The phonetic realization of the sentence is very faithful to its phonological organization as each tone is confined within syllable boundaries*

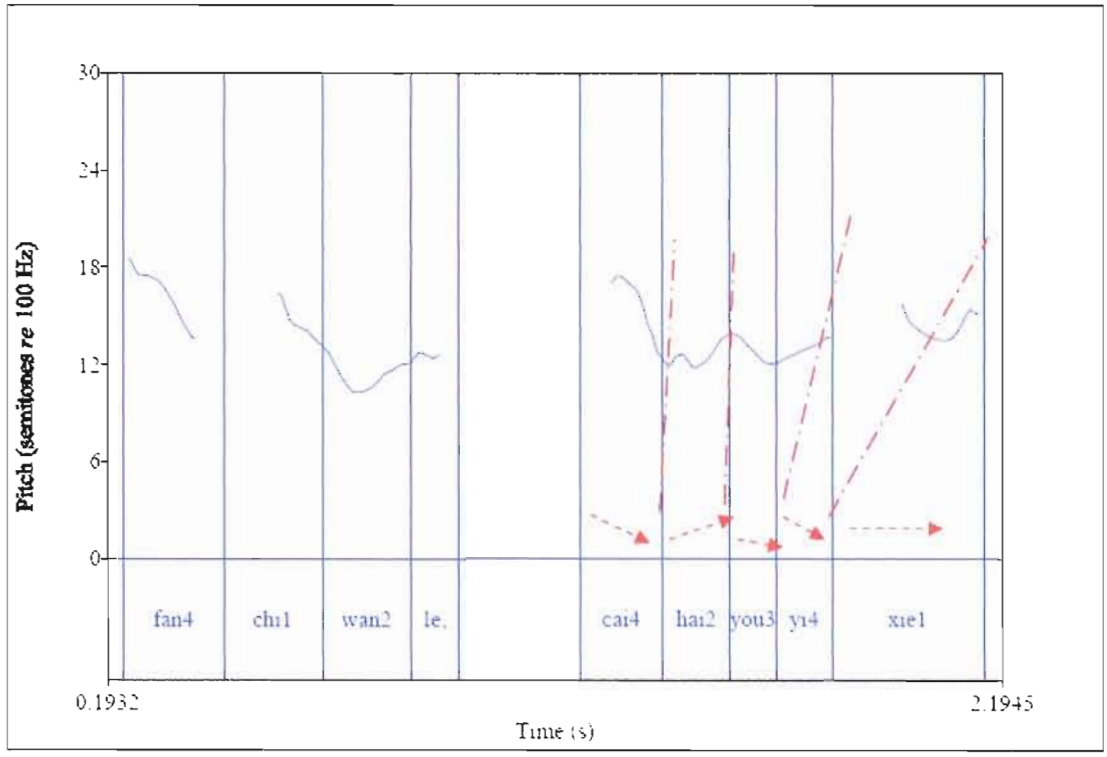


Figure 4.48: *Delay of syllables after tone 2 with a knock-on effect, that is to say, shift of all subsequent tones to the right. Sentence produced by one of the SC speakers taking part in the experiments.*

Despite these examples of peak delay in SC, phonetic alignment must be closer to phonological association with stressed syllables in SC than it is in English for the following reasons: peak delay in English can occur with any stressed syllable and as the consequences are very limited it takes place quite frequently. In SC however, peak delays only occur in specific contexts and can have semantic consequences.

Moreover, Chen (2007) found that in the speech of SC subjects producing Spanish words "the boundary of the rising and the following tone falls exactly on the syllable boundary [which] suggests that the Chinese learners tend to associate the rising and falling pitch contour with two separate syllables." The same author also found that SC subjects "tend to associate the rising pitch contour with the stressed syllable".

This pattern was particularly clear when the stressed syllable was preceded by an unstressed syllable.

The hypothesis is thus that SC speakers tend to align pitch accents with stressed syllables more accurately than NAE speakers do. One might further hypothesize that the contour of stressed syllables, likely to be a rising contour, will be produced within syllable boundary but might influence pitch alignment of the next syllables because of physiological limitations in adjusting articulatory gestures.

4.5.4.4 Results

Looking at the production of English by native speakers, we can see that the phonological peaks may be quite faithfully aligned phonetically or may be delayed (figure 4.49). It is also possible to have early peaks as in Figure 4.50 where the peaks of "old" and "man" are almost combined to produce one peak over the boundary.

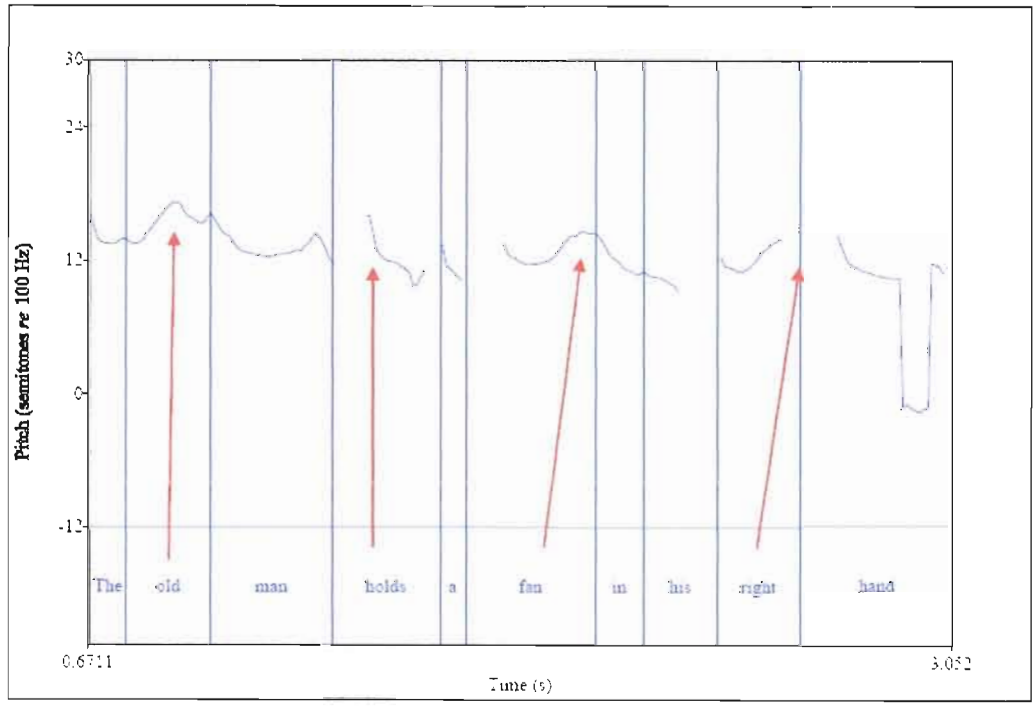


Figure 4.49: *Examples of peak alignment ("old" and "holds") and peak delay ("fan" and "right") by native speaker of English.*

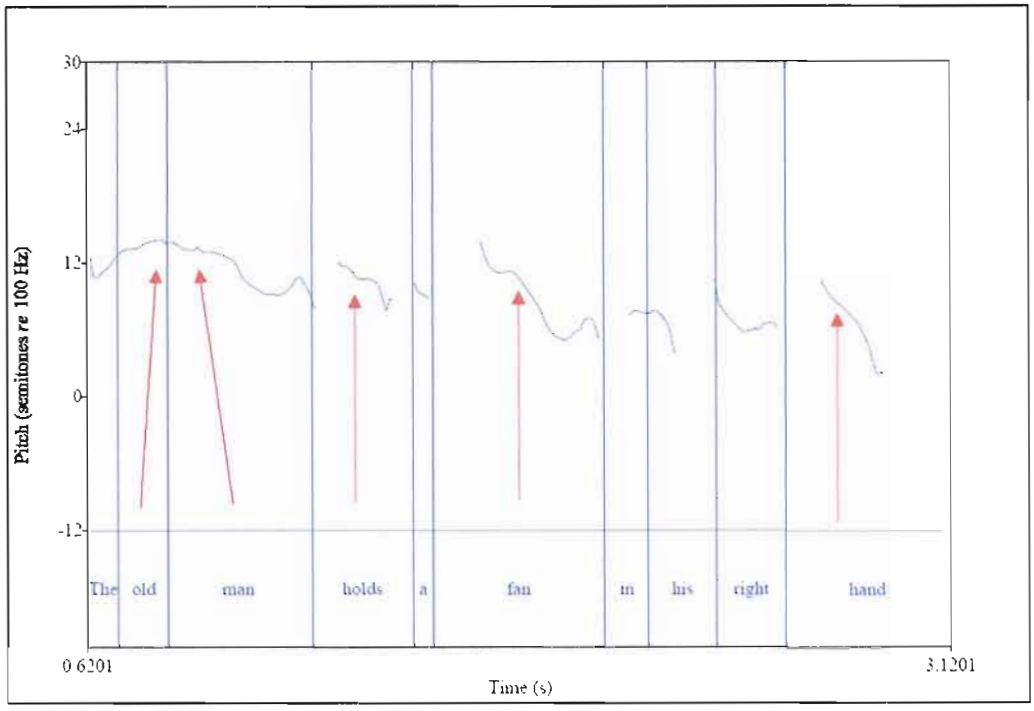


Figure 4.50: *Examples of early peak alignment ("man") very close to the peak of "old" by native speaker of English..*

We can also observe that transition from one peak to the next is quite smooth. For instance, the contour of "in his" is a natural continuation of the "fan" contour. Some of this can also be seen in the Chinese data, notably with the contour over "dìng yào" where the second 4th tone starts off where the first stopped. If one holds that a tone starts at a particular pitch level, one should expect the second tone to be a copy of the first one, in terms of its F0 features such as its beginning and end pitch levels, as represented in figure 4.51a. What we see however is that the second 4th tone is a continuation of the first, as schematized in figure 4.51b.

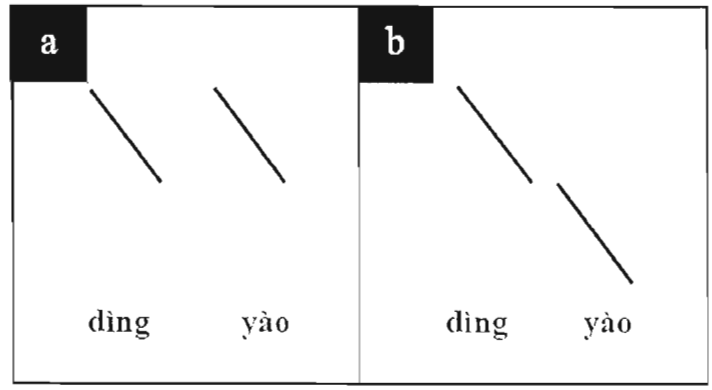


Figure 4.51: *(a) theoretical repetition of fourth tone in which both tones would start at the same level (b) actual repetition of fourth tone where the second tone starts where the first ends.*

This phenomenon of linking the production of tones or pitch accent will here and after be called "prosodic fluency". This term has also been used by Gu and Wan (2004) who give it the following description:

> *Prosodic fluency is closely related to naturalness. It evaluates continuity of prosodic characteristics within and between syllables. For examples, sudden change of pitch-contour height or intensity between two adjacent syllables will be perceived as prosodic discontinuity.*[44]

The difference between peak alignment and prosodic fluency must be reiterated: peak alignment is about considering whether of not the pitch gesture is completed within syllable boundaries; prosodic fluency is about joining, or linking, prosodic contours so as to avoid F0 discontinuity.

[44] These authors point out the distinction between prosodic fluency and acoustic fluency. The latter evaluates continuity of acoustic characteristics such as formant-traces at syllable boundary. This distinction is acknowledged and adopted here and only prosodic fluency is analyzed.

Looking at the production of the English sentence by SC speakers[45], we observe that some speakers consistently keep the F0 contour of each syllable confined to the syllable boundaries. For instance, this is clearly visible in the production of SC11 whose F0 contour is presented in figure 4.52. Other such examples, SC03, SC17, SC18, SC19, SC21, SC01, corroborate this observation. We can see that each contour is completed within the syllable boundaries. This production also lacks prosodic fluency as the individual contours do not appear as natural continuation of the preceding contour.

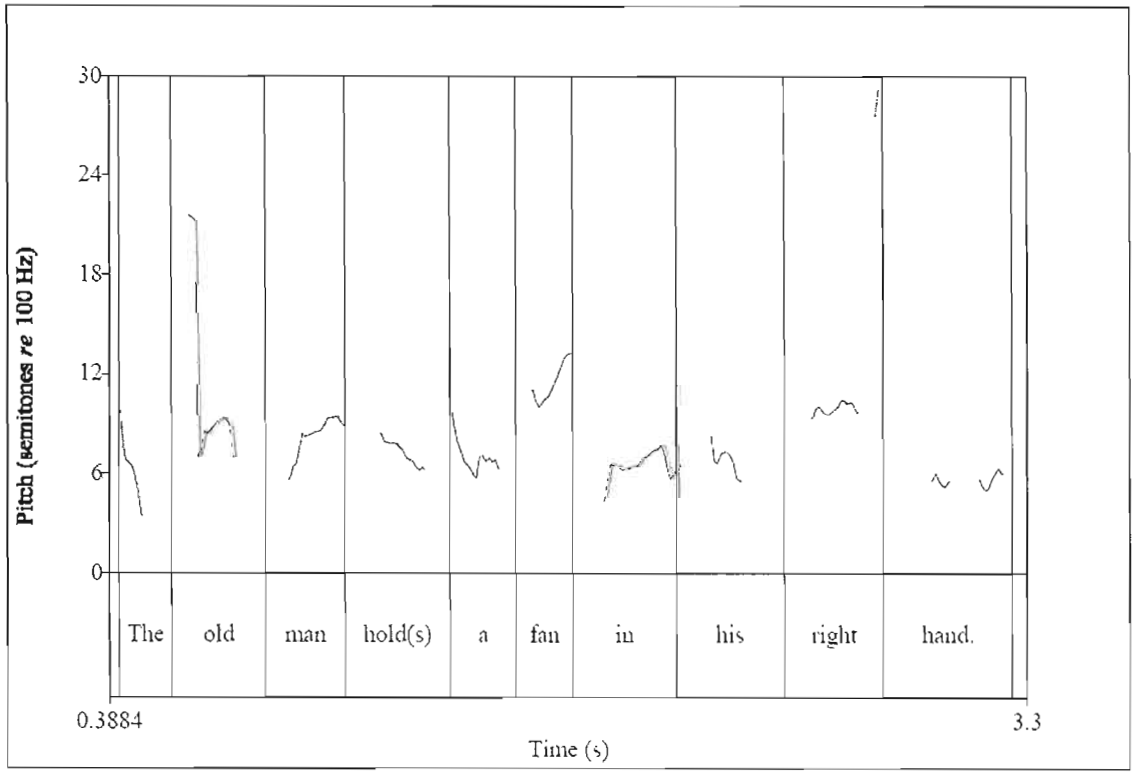


Figure 4.52: *Production of English sentence by SC speaker in which the F0 contour of each syllable confined to the syllable boundaries.*

Perceptually, this production is interpreted as particularly choppy. It must be noted though that the speaker did not produce pauses like one might expect of very low

[45] The complete inventory of SC speakers' pitch contour segmentation is available in order of ranking in Appendix 4.

level students. The speaker's production is clearly that of a sentence rather than that of a list of syllables; in other words, the subject did not produce the words as if they were shown sequentially, i.e one after the one.

As was expected, SC speakers produced the stressed syllables, except the final syllable, as rising contours. A very clear example of this is provided in figure 4.53. Again, we see that all pitch accents are produced within the phonological boundaries of the syllable[46]. We also see that the syllable after a high rise includes a fall before the next pitch accent which, in this case, is another rise. However, contrary to the SC example and to the production by native speakers of AE, this second pitch accent is compressed so as to fit within syllable boundaries; its peak is not delayed.

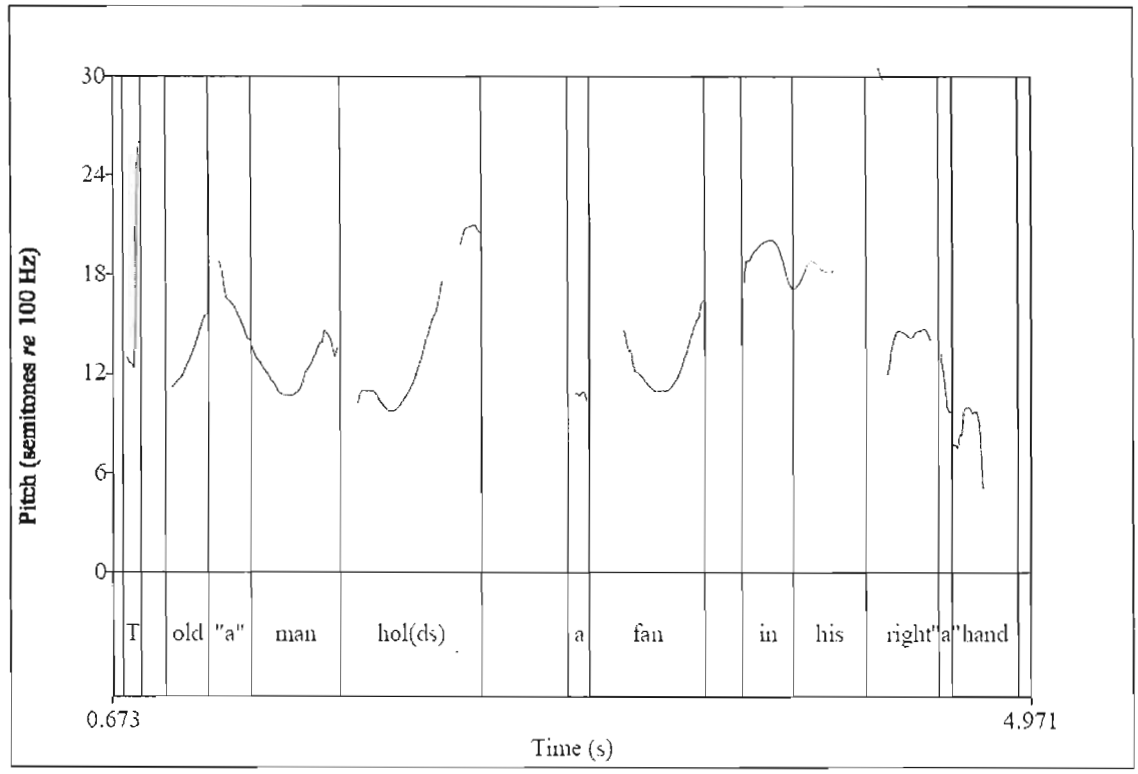


Figure 4.53: *Production of stressed syllables as high rising tones on "old", "holds" and "fan" by SC speaker.*

[46] The peak on "in" must not be interpreted as the end of the « fan » peak. Indeed, for unknown reasons, the speaker emphasized the preposition, which gives it this particularly high F0. The 175ms pause between the two words substantiates that the two contours are unrelated.

While the subjects all produced the pitch accents within syllable boundaries, some of them show signs that peak delay would be possible.

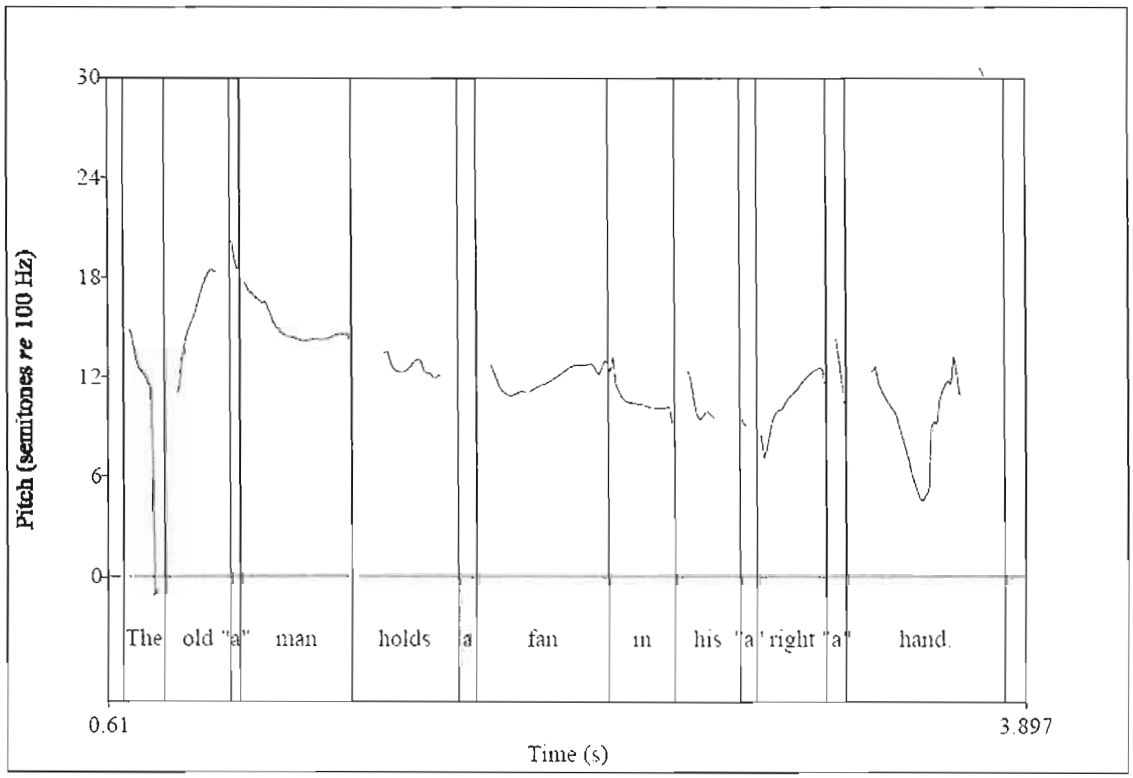


Figure 4.54: *Example of contour peak on "fan" reached before the end boundary.*

For instance, in figure 4.54 above, we can see that the contour of "fan" reaches its peak before the boundary. However, we also see that the beginning of the contour of the next word is quite a smooth and natural continuation of the "fan" contour. This is what was earlier described as prosodic fluency. This smooth connection of contours would seem necessary to produce peak delay because it sets up perfect conditions for the peak to glide off to the next syllable. If the transition between pitch accents is very abrupt, as in figure 4.53 ("old man"), peak delay could not occur without prosodic error. Indeed, if a high peak is delayed, the first syllable might appear unstressed because of the great difference of pitch between the early part and the end of the contour. This is illustrated in figure 4.55 where, in 4.55a,

syllable 1 is stressed although the peak is on the boundary between the two syllables. In figure 4.55b however, the slight delay of the peak, transforms the first syllable into an unstressed syllable and the second as a stressed syllable, a scenario contrary to the situation in figure 4.55a.

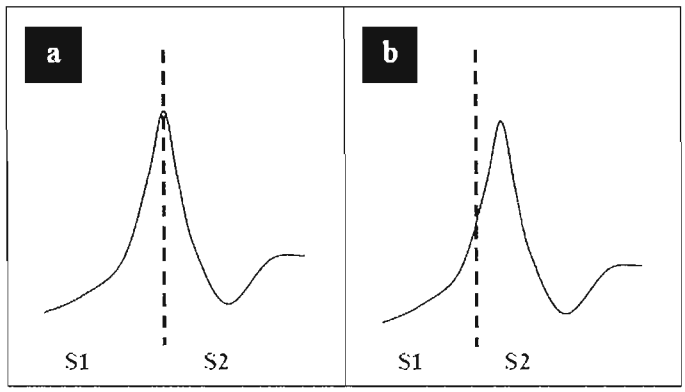


Figure 4.55: *Schematized peak delay of high rising contour causing perception of stress shift. In a) the stress is perceived on syllable 1 (S1) although the peak is on the boundary between S1 and S2. in b) the stress is likely to perceived on S2 because S1 has a very low rising contour typical of an unstressed syllable.*

The five speakers (SC13, SC07, SC16, SC25 and SC14) who produced greater prosodic fluency also produced sentence contours that closely resemble those of native speakers, which could be schematized as three rounded peaks, as illustrated in figure 4.56.

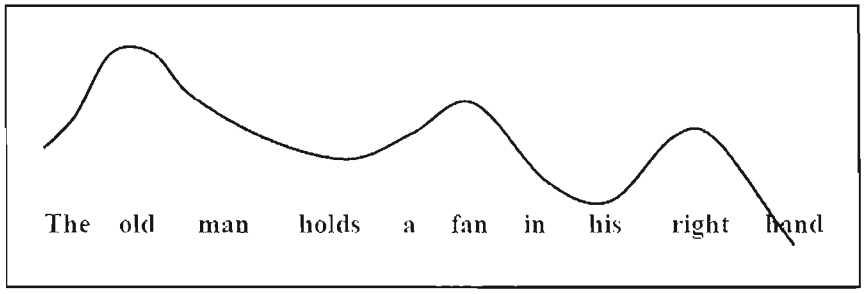


Figure 4.56: *Schematized prosody of the sentence produced with three rounded peaks over "old", "fan" and "right" as produced by native speakers of English and by the more fluent sounding SC speakers.*

Perceptually, these productions are interpreted as less choppy than sentences in which the joining of contours is more abrupt.

These speakers were respectively ranked 7th, 14th, 15th, 16th, and 21st. It thus shows that while smoothness might have played a part in the evaluators' judgement of prosodic competence, other factors weighted heavily as well.

4.5.4.5 Conclusion

This experiment has shown that the SC speakers align pitch accents with stressed syllables more accurately than NAE speakers do. The research also confirmed Chen's finding that SC speakers associate rising pitch contours with stressed syllables. High rising pitch contours were also found to have an impact on the syllable immediately following the stressed syllable but, unlike in SC, not the subsequent ones.

CHAPTER V

DISCUSSION AND CONCLUSION

5.1. Introduction

The stated aim of this study is to bridge research on phonological investigation with second language research. Throughout this work, phonology and phonological organization have been investigated from a linguistics perspective with an outlook on ESL acquisition and on the production of ESL prosody. This final chapter consolidates the connection between the two domains in difference ways, notably by exposing the implications of the results obtained from the four experiments and providing potential applications in the classroom.

5.2. Summary of results and discussion

In this thesis, two aspects of phonological theory were investigated: typology and phonological organization. This choice was driven by the common belief that prosodic foreignness in a second language is due to one of two factors: a difference of typological class of L1 and L2; a transfer of L1 prosodic features.

The study of typology led to the conclusion that, at this stage, no prosodic typology is valid, at least not for application to L2 acquisition. More specifically, it was clearly shown that some typologies, particularly Pike's stress-time and syllable-time theory, should be definitively dismissed. The strength and assertiveness of this statement reflect not only the fact that this typology does not provide answers but also that it hinders progress in a number of fields and most specifically in second language acquisition. Indeed, this theory, which still prevails despite having been experimentally discredited, has conducted many to use it as an easy explanation for second language prosodic errors and problems. Because the stress-time syllable-

time theory provides an erroneous explanation to learners' errors, there is little hope of finding solutions to correct these errors.

The conclusion of the typology section of this thesis was thus that no typology can be used to explain second language prosody. Some of the typologies were however shown to have potential application in other fields such as Low and Grabe's (1995) Pairwise Variability Index for speaker variation within a language. The only exception made was the potential application to second language acquisition of Speech Cycling (Cummins, 1997) not as a theory or explanation but as a pedagogical tool which could bring learners to adopt a native speakers' "beat". This would basically consist in prosodic drills: repetition of phrases and utterances for their prosodic properties. Whether, upon acquiring the correct speech cycle, learners would be able to transfer their skill to other segments with similar or different prosodic patterns remains to be proven. Speech cycling could thus become part of a pedagogical methodology.

The second aspect of phonological theory investigated in this thesis is phonological organization. This was conducted to test the likelihood of L1 to L2 transfer of phonological organization of prosody. The analyses of North American English, French and Standard Chinese yielded important organizational differences between North American English and the other two languages. This was used to posit hypotheses about ESL prosody by native speakers of French and Standard Chinese. While such research is a comparative analysis, it should not be mistaken for the application of Lado's (1957) Comparative Analysis Hypothesis. The latter claimed that comparison of surface features of two languages permitted to anticipate all positive and negative transfers between those languages; this categorical and uncompromising theory has long been experimentally discredited. In fact, no specific acquisition model was adopted for this thesis. Only the generally acknowledged idea of L1 to L2 transfer was assumed. However, in this thesis, the

premise is that differences at the underlying prosodic organization level rather than at the surface feature level are transferred.

The conclusions of the experiments are respectively summarized hereafter:

- Experiment 1 tested FF and QF speakers' ability at producing rhythmically simple sentences. As hypothesized, the subjects produced adequate prosody except for early AP function words that were stressed while they should not have been. Contrary to hypothesis, end of word-group (i.e. AP) pitch rise was not more a characteristic of FF ESL prosody than of QF ESL prosody. In fact, in some sentences, more QF speakers produced rises than FF speakers. Furthermore the level of occurrence of these rises (for sentences 1 and 2 respectively, 4 and 1 out of 11 FF and 4 and 5 out of 20 QF speakers) does not permit to generalize that end of AP rises are transferred from French L1. Moreover, these occurrences did not corroborate with speaker ranking while other prosodic features did. Indeed, incorrect phonological stress, slow rate of speech, use of pauses and increased length of these pauses all correspond with lower-ranked speakers.

- Experiment 2 tested FF and QF speakers' ability at producing rhythmically more difficult sentences. We saw that basic stress placement rules had been assimilated by the majority of speakers. The more difficult (i.e. less frequently used) rules and non-rule based stress were assimilated by only a few speakers who were not necessarily high-ranked. The same comment is valid for AP final syllables which the vast majority of speakers stressed. An unexpected finding of this experiment is that final syllable rises do not affect native speakers' perception of stress; that is to say that these rises were not perceived as stress. The relative difficulty judges had in locating the stressed syllable in polysyllabic words was also surprising. It was explained from the experimental analysis which showed that different stress correlates were attributed to different syllables. Indeed, while native speakers produced

greater F0, intensity and duration on the same syllable – the stressed syllable –, FF and QF speakers could attribute these correlates to different syllables. This led to the observation that perception of stress by native speakers of English seems to result from a matrix of relative importance of these correlates and of energy and pitch contour. The experiment also highlighted the fact that perception and instrumental analyses can provide divergent information and that thorough research should combine experimental methods.

- Experiment 3 analyzed SC speakers' production of English pitch accents and SC tones. The hypothesis was that SC speakers would produce the SC tone that most closely matched the pitch contour needed in English. Experimental analysis of F0 contours showed that SC speakers indeed used the contour and pitch range of tone 2 to produce all rising pitch accents. This was first observed with combined data and confirmed by individuals' tone 2 idiosyncrasies as there were also visible in English pitch accents.
  We saw also that the vast majority of speakers used tone 2 (a rising contour) at the end of a phrase that introduces a list where native speakers of English typically use a high falling tone. This does not discredit the "tone for pitch contour" hypothesis but rather indicates that the prosodic features of this type of phrase boundary have not been assimilated by SC speakers.

- Experiment 4 analyzed SC speakers' alignment of pitch accents with stressed syllables. The hypothesis was that, as tones are generally aligned with the segmental elements of the syllable, SC speakers would align pitch contours more neatly than NAE speakers do. The secondary hypothesis, that SC speakers would associate rising pitch contours with stressed syllables, was confirmed but also permitted to suggest a correspondence tone contour to peak delay. Indeed, it was shown that the transformation of a rising pitch contour from very sharp to more curved is necessary for the production of peak delay.

5.3.Pedagogical Interpretation of results

To be useful to second language acquisition of prosody, these findings must be interpreted with a different perspective, i.e. the nature of the errors made by the subjects; this is done hereafter.

5.3.1. Nature of errors

Lexical stress, that is to say, whether or not a monosyllabic word should be stressed and which syllable in polysyllabic words should be stressed, is generally assimilated by FF and QF speakers. In figure 5.1, all lexical stress errors are marked by cones of different shades of blue. We should first bear in mind that many words do not appear; in fact, only 8 of the 35 words used in the 5 sentences were incorrectly stressed by one or more students. Among those, only three words (in light blue) caused problems to a fair proportion of the subjects. We can see that those lexical errors were not consistently produced by lower ranked subjects.

Phonological prosodic stress however caused more trouble for the same speakers. The areas that were identified as potential negative transfer sites did in fact prove to be problematic for these subjects. Notably, we saw that what francophone speakers identify as accentual phrases (AP) were prosodically marked in English the way they would be in French, that is to say, with an early and a late rise. These errors were quite consistently produced by lower-ranked subjects. We can also see that the AP final stress ("him") was more of a problem than early AP stress ("on" and "in"). Indeed, "him" was stressed more strongly and by more speakers than the other two words. We also saw that polysyllabic words often were produced with a final rise.

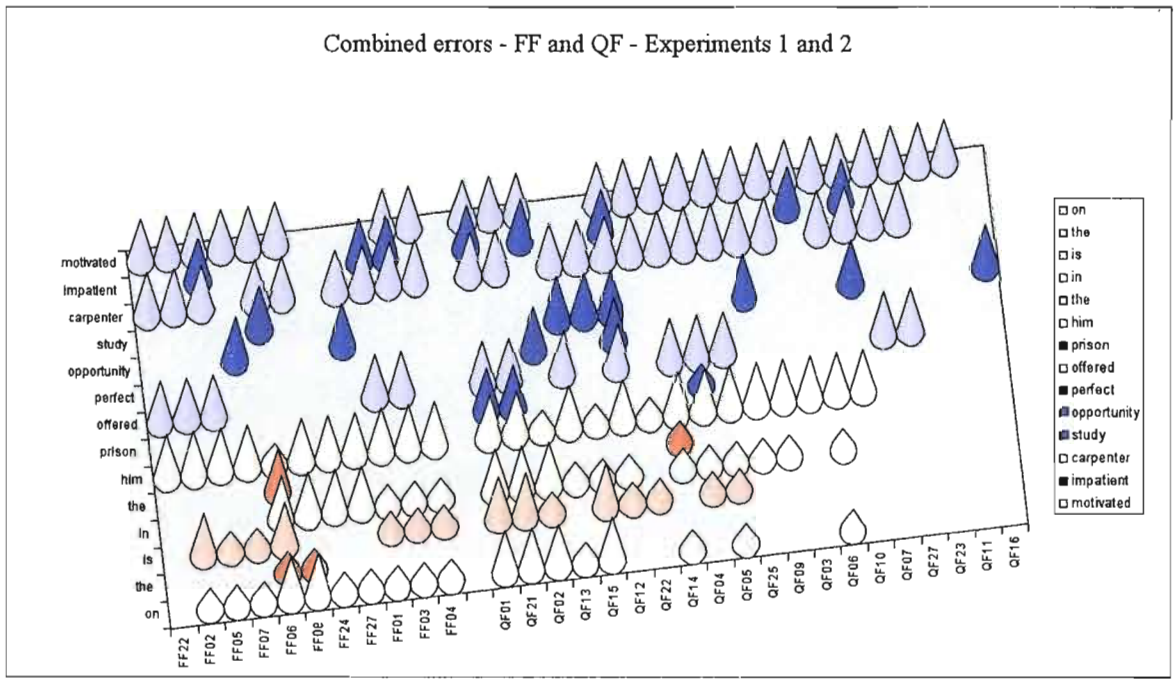


Figure 5.1: *Diagram of all stress errors produced by FF and QF subjects in experiments 2 and 3. Lexical errors are indicated by blue cones and phonological by yellow/orange cones). Each cone represents a stress error: presence of stress, incorrect placement, multiple stressed syllables.*

Phonetic correlates of prosody such as slow rate of speech, use of pauses, increased pause length were also seen to concur with low rank of competence of francophone subjects. This is particularly evident from the simple sentences of experiment 1. Because the correlation between these features and rank is so clear, it is apparent that rate and pauses are decisive in the judgment of level of competence. However, as the judges did not report any problem understanding the subjects, these phonetic features do not seem to affect comprehension but only the perception of fluency.

In the case of polysyllabic words, the incorrect use of phonetic correlates was felt as a hindrance to comprehension. Indeed, the allocation of the stress correlates to different syllables thwarted the understanding of the stress pattern and of the word's

segmentation. The sentences used were quite lexically and syntactically simple but one can easily speculate that longer and more complex sentences could be an issue for listeners. While we saw that word final rises were not considered to be stressed by the judges, we also noted that they interfered with perception of fluency.

Phonetic correlates were also shown to be transferred in the case of SC speakers. This was attested in experiment 3 with the use of tone contours for pitch contours. As SC tones are a lexical feature, this indicates that SC speakers transferred phonetic correlates from lexicon to prosody and/or that the speakers attributed prosodic features to syllables rather than to prosodic domains, i.e. levels of phrasing.

This second proposal was supported by the findings of experiment 4 which showed that some SC speakers consistently kept the F0 contour of each syllable confined to the syllable boundaries. Moreover, as we saw that in SC tone and stress are inseparable, we can expect that subjects who attribute tone to individual syllables will have the tendency to stress all syllables. Figure 57 for instance shows that all syllables are quite equally stressed by the speaker.

5.3.2. Ranking

For the experiments presented in this thesis, the speakers were ranked in order of competence. This was necessary because it is assumed that transfers are more frequent and more serious in beginners' speech than in advanced learners' speech. By taking this factor into account, we obtained different pictures of L2 prosodic production at different stages of interlanguage (the series of progressive levels of competence reached by non-native speakers) for FF and QF speakers. If this factor had not been taken into account, only one picture, much less precise, would have emerged. From the representations we did obtain, we catch a glimpse of the progression that speakers of a same language are likely to sustain in their production of English prosody. A complete extrapolation is not possible because the same

number of subjects at each level of competence would be needed in order to have more reliable data. However, according to the judges' evaluation, the subjects were quite evenly distributed along the scale (see "number of points" in the three tables in Appendix 6) and some conclusions can be suggested. For instance, we can see that, within phonological prosodic stress, subjects who no longer produce early AP rises continue to produce final AP stress. We also saw that lexical stress is assimilated more quickly than prosodic stress. Second language professionals (textbook writers, pedagogical method developers and teachers) could therefore be well advised to take these points into consideration when developing new material. Indeed, the features that seem to be assimilated quickly and, possibly, fairly naturally (i.e. without recourse to substantial pedagogical support), arguably should not be the pedagogical focus of prosody teaching. Understanding the likely progression in prosodic competence from a given L1 should also be of help to L2 academic professionals. In the example of SC production of rising pitch accents, we saw that while subjects produce sharply rising peaks, peak delay is unlikely. It is not suggested here that delayed peaks are important to perception of prosodic competence or of fluency and that ESL should learn to produce them. Peak delays are however characteristic of a certain level of fluency. It should therefore be considered that getting Chinese ESL learners to produce rising peaks that are more rounded would help these subjects to improve or enhance their production of prosody.

In other cases, we saw that a same error was produced by a high proportion of the subjects but not necessarily low-ranked subjects. For instance, we saw that many SC speakers attributed an F0 contour to individual syllables. As explained above, this indicates that, for these speakers, F0 variation is constrained predominantly by the lowest domain in prosodic hierarchy, the prosodic word (PW), rather than by higher level domains: the intonational phrase domain (IP) or the intermediate phrase (ip) domain. Support of the latter proposal comes from the observation that SC speakers produced all "fan" words with the same type of contour, a rising contour,

i.e. a tone 2. One of those words, the first occurrence of "fan" should however have been pronounced with a falling tone as required by sentential prosody. This implies that SC speakers did not recognize the phrase as a prosodic domain with its own boundary tone and thus applied a syllabic tone a rising tone, as by default.

5.3.3. Listeners' perception

One of the aspects that must be borne in mind with pronunciation training is that it targets the learner for the benefit of the listener, and subsequently, for the benefit of this learner. Indeed, improved pronunciation helps the listener to more easily understand the speaker. After all, generally speaking, with some effort, any native speaker can understand a non-native speaker[47]. Books and course titles that include the term "accent reduction", "accent neutralization" or "corrective phonetics" make this listener perspective of pronunciation very clear. To emulate a popular saying: "Accent is in the ear of the hearer". Pronunciation training must therefore take the *listeners* into account because *their* perception of foreignness is what should direct book and program design. This is why the correlation between judges' evaluation and L2 prosodic features is crucial to designing efficient pronunciation training. We saw for instance that use of pause and rate of speech clearly correlate with low rank; we can therefore assume that these affect the listener's perception of level of foreignness. With regard to polysyllabic words, the judges reported that final syllable rises were distracting and that the use of different stress correlates (or phonetic features such as duration, energy, intensity) on an array of syllables hindered comprehension and made it difficult to segment speech. These comments are crucial and should be the basis of pronunciation training. Native speaker judgment should also be used to distinguish which prosodic errors affect comprehension and which affect the perception of foreignness.

[47] Admittedly, in some rare cases, speakers are close to incomprehensible; often times, issues other than pronunciation skills (speech impediment, extremely limited exposure to the spoken form of the language) are in play. Equally, some people are little receptive to any accented form of their language; a native speaker of American English might for instance find it difficult to understand Spanish- or French- or British-accented English.

5.4.Pedagogical application

In pedagogical terms, these data translate into a necessity to bring these speakers to learn to produce words in prosodic context: in groups and in various sentential positions. That is to say, the focus should be moved away from phonemes, syllables, and words and placed on the prosodically adequate production of word groups. Besides, teachers have long been aware that learner's production of phonemes in isolation is in no way a guarantee of the production of that same phoneme in a word or in a group of words. Similarly, polysyllabic words that are adequately stressed in isolation are often incorrectly stressed in a sentence. Of course, ESL learners need to practice phonemes and syllabic issues (such as consonant clusters) but this can also be done within a prosodic framework. For instance, if we take the example of verbs in "-ate" and their derivatives, extensive exercises could be designed with a variety of such words placed in different positions of sentences of varying length and of various phonological contours. The focus would remain the prosody of the sentence, which includes the adequate stressing of the –ate word. In the case of SC speakers, it would be recommended to ensure that the target word be placed just before a phrase boundary for which a falling tone (L-) is required so that the subject dissociates phrasal stress and rising tone.

5.5. Pedagogical implications

5.5.1. Consideration of L1 prosody

The experiments of this thesis have shown that L1 prosodic features are transferred in a number of ways, in a variety of levels of perceived gravity and with a number of consequences on native speakers' perception of foreignness.

One of the direct implications of these findings is that, for efficient training, L1 prosodic organization must be taken into account in the teaching of second language prosody. This certainly has great implications as it means that, for many languages, L1 phonological organization will need to be developed. As we saw, much work has been done on English and French but much less has been written about Chinese

prosodic organization, even if interest in the latter has grown in the last decade. The prosody of many languages has been very little researched, often because the labeling of these languages as "stress-timed" or "syllable-timed" was assumed to be sufficient explanation. Once the prosodic phonology is understood, the comparison of the two languages will be needed and the differences evaluated.

In this thesis, no notable and consistent difference between FF speakers' and QF speakers' production of ESL prosody was found (experiments 1 and 2). Prosodic differences are however attested in language varieties. While the comparison between varieties need not be the focus of L2 prosody acquisition, the selection of a variety over another would have to be carefully thought through.

5.5.2. Classroom composition

If one admits that the learners' L1 is important, one must deal with the issue of an ESL clientele from a diversity of linguistic backgrounds. This is a serious consideration for academic institutions and programs that cater to local natives, exchange students and immigrants[48]. In such context, a teacher can be faced with a classroom where more than ten native languages are spoken. Teachers who have had this experience know how difficult it is to introduce material critical to one part of the group but of no use to the rest. While the use of laboratories can help in the management side of providing material to particular language groups, specific material still needs to be developed.

The situation is usually less problematic in institutions where English is taught as a foreign language (EFL) and where students share a common language. Nonetheless, attention still needs to be paid to the students' linguistic background as some of them

[48] This issue is not unique to the teaching of prosody but of pronunciation and of language in general; it does however remain an issue.

may speak another language at home, as is often the case of immigrants and their children.

This issue also concerns textbook writers whose work is not directed at one particular L1 group. Labeling exercises for speakers of particular languages, as Dauer (1993) did for phoneme practices, is a possibility.

We saw earlier that aspects of prosody would need to be assimilated in a particular sequence. That is to say, that some aspects might be learnt before others because they offer a springboard to understanding these other aspects. In other words, although two groups of different linguistic backgrounds would need, for example, to learn to produce more rounded peaks, it could be a priority for one group but better dealt with later for the other group.

This additional difficulty could only be overcome with modular training, i.e. the creation of an individualized syllabus for each student. Financial and practical considerations are likely to make this proposition idealistic.

5.6. Limitations of this thesis and directions for further research

In this study, a number of aspects of ESL prosody were analyzed and four experiments were conducted. While it provides some answers with regard to English prosodic features of ESL by speakers of French and of Standard Chinese, it does not provide a complete picture. For instance, more experiments would need to be conducted to test all of the propositions (and possibly others) that were made following the comparison of the phonologies. This thesis suggested that some features might be (or need to be) assimilated before others. It did not however prove this conclusively and further research is needed in this field. The very conclusions of this thesis generate limitations of its validity: for instance, we saw that different levels of proficiency were characterized by certain features. For greater validity, a

greater number of subjects at each level would be required. We also saw that, for FF and QF speakers, some features were closely related to the perception of competence by native speakers. More needs to be understood about this, as improving pronunciation is about improving communication. Furthermore, classification of which features impede understanding and which are perceived as foreign accent would be extremely useful, especially for pedagogical applications. Finally, the experiments conducted with these language groups could be conducted with other language groups. The choice of doing different experiments for the two language groups was deliberate and justified in 4.1. Conducting the same experiments with other language groups would enable the creation of a platform to compare ESL production and to confirm whether or not the errors are language specific.

5.7. Concluding remarks

The aim of this thesis was to bridge research on phonological investigation with second language research and more specifically, to conduct an investigation of prosody whose results could serve the needs of teachers who deal with the acquisition of prosody by learners of English as a second language.

It indeed showed that phonological research can serve the needs of second language acquisition research. Like any efficient bridge, this one is bidirectional. This means that teachers can provide important feedback on the application of the research. For instance, teachers can observe the feasibility and benefits of teaching certain features and answer a number of questions: are all features learnable? Does their acquisition affect (improve or worsen) other features? Are some subject impervious to some features and their acquisition? Peperkamp and Dupoux (2002) talked about "stress deafness" for instance. Theoreticians would be able to interpret these answers according to their model of choice, for instance in terms of universality or markedness.

No particular model was adopted for this thesis but it assumed comparison of L1 and L2 to be important because it supported the concept of prosodic transfer. The research conducted here successfully demonstrated the validity of this assumption, including that of tonal transfer.

BIBLIOGRAPHY

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press acoustic study.

Allen, G. 1975. "Speech rhythm: Its relation to performance universals and articulatory timing". *Journal of Phonetics,* 3, pp. 75-86.

Anderson-Hsieh, J., J. R. Johnson and K. Koehler. 1992. "The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure". *Language Learning*, 42, pp.529-555.

Arvaniti, A., and G. Garding. 2007. "Dialectal variation in the rising accents of American English". In Cole, J. and J. Hualde (eds.), *Laboratory Phonology 9*. Berlin and New York: Mouton de Gruyter. http://idiom.ucsd.edu/~arvaniti/Arvaniti&Garding.inpress..pdf

Auer, P. 1993. "Is a rhythm-based typology possible? A study of the role of prosody in phonological typology". *KontRI Working Paper No. 21.* Konstanz: Fachgruppe Sprachwissenschaft, Universität Konstanz. http://www.germanistik.uni-freiburg.de/auer/?download=Phonotypo_Kontri1.pdf

Balusu, R. 2001. "Acoustic correlates of stress and sccent in Telugu". Paper accepted at *SALA XXI*. Konstanz, Germany 2001. http://homepages.nyu.edu/~rb964/correlates.pdf

Beckman, M., and J. Pierrehumbert. 1986. "Intonational structure in Japanese and English". *Phonology Yearbook III*, pp. 15-70.

Beckman, M., and J. Pierrehumbert. 1986. "Intonational structure in Japanese and English". *Phonology Yearbook III*, pp. 15-70.

Bertrán, A. P. 1999. "Prosodic typology: on the dichotomy between stress-timed and syllable-timed languages". *Language Design,* 2, pp. 103-131

http://elies.rediris.es/Language_Design/LD2/pamies.pdf#search=%22Prosodic%20Typology%3A%20On%20the%20Dichotomy%20%22

Bissonnette, S. 1997. "Comparaison du registre de locuteurs québécois et de locuteurs français ». In *Actes des 11èmes Journées de Linguistique.* Québec : CIRAL, Université Laval, pp. 17-25.

Boersma, P., and D, Weenink. 2007. *Praat: doing phonetics by computer* (Version 4.5.17) [Computer program] Retrieved March 21, 2007. http://www.fon.hum.uva.nl/praat/

Boula de Mareuil, P., and Vieru-Dimulescu, B. 2006. "The contribution of prosody to the perception of foreign accent". *Phonetica*, 63, pp. 247-267.

ftp://tlp.limsi.fr/public/Boula_Vieru_phonetica06.pdf

Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective.* Lund: Gleerup.

Cao, J., Lu, S., and Y. Yang. 2000. "Chinese prosody and a proposed phonetic model". *Chinese Academy of Social Sciences Report of Phonetic Research 2000.* pp. 27-32 http://ling.cass.cn/yuyin/report/files/2000_5.pdf

Cauldwell, R. 2002. "The Functional irrhythmicality of spontaneous speech: A Discourse view of speech rhythms". *Apples 2/1*, pp. 1-24

http://www.solki.jyu.fi/apples/020201/The%20functional%20irrhythmicality%20of%20spontaneous%20speech.htm

Cedergren, H. J., H. Perreault, F. Poiré and P. Rousseau. 1989. " L'accentuation en français québécois : implémentation d'une approche tonale", *Revue Québécoise de Linguistique*, 19:2, pp. 19-32.

Celce-Murcia, M., D.M. Brinton and J. Goodwin. 1996. *Teaching Pronunciation: A Reference for Teachers of English as a Second or Foreign Language.* Cambridge, MA: Cambridge University Press.

Chan, M. 1993. "Review of the prosody of Mandarin Chinese". *Journal of Phonetics*, 21:3, pp. 343-347.

http://people.cohums.ohio-state.edu/chan9/pubn/shen-rev.htm

Chang, N. 1958. "Tones and intonation in the Chengtu dialect (Szechuan, China)". *Phonetica*, 1-2:59-85. Reprinted in *Intonation*. Ed. by Bolinger. 1972. Great Britain: Hazell Waston and Viney Ltd. pp. 391-413. (Cited in Shen, 1990)

Chao, Y-R 1930. "A system of tone-letters". *Le maître phonétique,* 3rd series, 45. pp. 24-27.

_____ . 1933. "Tone and intonation in Chinese". *Bulletin of the Institute of History and Philology,* 4. pp. 121-134.

_____ . 1968. *A grammar of Spoken Chinese.* Berkeley: University of California Press.

Chen, G-T. 1974. "The pitch range of English and Chinese speakers," *Journal of Chinese Linguistics.* 2. pp. 159–171. (Cited in Wang et al. 2003.)

Chen, Y. 2007. "From tone to accent: the tonal transfer strategy for Chinese L2 learners of Spanish". *Proceedings of 16th International Congress of Phonetic Sciences (ICPhS XVI)*, ID 1043, Saarbrücken, 6-10 August 2007. http://www.icphs2007.de/conference/Papers/1043/1043.pdf

Chiang, W-Y., and F-M. Chiang. 2005. "Saisiyat as a pitch accent language: evidence from acoustic study of words". *Oceanic Linguistics*, 44:2 (December 2005) University of Hawai'i Press.

Collins Cobuild on CD-rom (2nd edition). 2002.HarperCollins Canada / UK Electro Dic

Coppieters, R. 1987. "Competence differences between native and near-native speakers". *Language,* 63:3, pp. 544-573.

Corder, S., 1967. "The significance of learner's errors". *International Review of Applied Linguistics in Language Teaching (IRAL),* 5, pp. 161-170.

Couper-Kuhlen, E. 1993. *English Speech Rhythm. Form and Function in Everyday Verbal Interaction.* Amsterdam: Benjamins.

Cruttenden, A. 1997. *Intonation*, 2nd ed. Cambridge: Cambridge University Press.

Crystal, D. 1969. *Prosodic Systems and Intonation in English*. London: Cambridge University Press.

Cummins, F. 1997. "Rhythmic coordination in English speech: an experimental study". Doctoral dissertation, Indiana, Indiana University.

_____ . 2002. "Speech Rhythm and Rhythmic Taxonomy". *Proceedings of Prosody 2002,* pp. 121-126, Aix en Provence.

http://aune.lpl.univ-aix.fr/sp2002/pdf/cummins.pdf

Cummins, F., and R. F. Port. 1998. "Rhythmic constraints on stress timing in English". *Journal of Phonetics, 26:2,* pp. 145-171.

Cutler, A. 1984. "Stress and accent in language production and understanding". in *Intonation, Accent and Rhythm: Studies in Discourse Phonology*, Gibbon, D., and H. Richter, (eds.), pp. 77-90. Berlin: Walter deGruyter,

http://books.google.ca/books?id=uqjKkXjzk9sC&pg=PA77&lpg=PA77&dq=%22stress+and+accent+in+language+production%22&source=web&ots=uN0vI6ScQ2&sig=AHlqxr92b298ETNIh6t5C6X1NkY&hl=fr#PPA80,M

Dauer, R. 1983. "Stress-timing and syllable-timing reanalyzed". *Journal of phonetics,* 11, pp. 51-62.

_____ . 1987. "Phonetic and Phonological Components of Language Rhythm". *Proceedings of the XIth International Congress of Phonetic Sciences*, pp. 268-274.

_____ . *Accurate English: A complete course in pronunciation.* Englewood Cliffs, N. J.: Prentice-Hall Regents.

DeFrancis, J. 1963. *Beginning Chinese.* Yale University Press. (Cited in Lee, 2005)

_____ . 1984. *The Chinese language: Fact and fantasy.* Honolulu: University of Hawaii Press.

Delattre, P. 1965. *Comparing the Phonetic Features of English, German, Spanish and French.* Heidelberg: Julius Groos.

Di Cristo, A. 2003. "De la métrique et du rythme de la parole ordinaire : l'exemple du français", *Semen*, 16, Rythme de la prose. [online]. http://semen.revues.org/document2944.html

Di Cristo, A., and D.J. Hirst, 1993. "Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français". *Travaux de l'Institut de Phonétique d'Aix*, 15, pp. 9-24.

D'Imperio, M. 2000. "The role of perception in defining tonal targets and their alignment". Doctoral Dissertation, Columbus, Ohio State University. http://linguistics.osu.edu/research/publications/dissertations/files/dimperio2000.pdf

Donegan, P., and D. Stampe, 1983. "Rhythm and the holistic organization of language structure". In *Papers from the parasession on the interplay of phonology, morphology & syntax* edited by F. Richardson et al. pp 337-353. Chicago Linguistics Society. (Cited in Auer, 1993)

http://www.ling.hawaii.edu/austroasiatic/AA/rhythm1983.rtf

Duanmu, S. 2000. *The phonology of Standard Chinese*. Oxford: Oxford University Press.

_____ . 2004. "Tone and non-tone languages: an alternative to language typology and parameters." *Language and Linguistics.* 5:4, pp. 891-923.

_____ . 2005. "The tone-syntax interface in Chinese: some recent controversies," *Proceedings of the Symposium "Cross-Linguistic Studies of Tonal Phenomena, Historical Development, Tone-Syntax Interface, and Descriptive Studies"*, edited by Shigeki Kaji. Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies, pp. 221-254. http://www-personal.umich.edu/~duanmu/ToneSyntax05.pdf

Eckman, F. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning,* 27, pp. 315-330.

Flege, J. E. 1995. "Second-language speech learning: theory, findings, and problems". In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange, pp. 229-273. Timonium, MD: York Press.

Fónagy, I., and J. Fónagy, 1976. "Prosodie professionnelle et changements prosodiques". *Le français moderne*, 44, pp. 193-228.

Fox, A. 2000. *Prosodic Features and Prosodic Structure. The Phonology of Suprasegmentals.* Oxford: Oxford University Press.

Fry, D.B. 1955. "Duration and intensity as physical correlates of linguistic stress". *Journal of the Acoustic Society of America,* 27, pp. 765-768.

_____ . 1958. "Experiments in the perception of stress". *Language and Speech,* 1, pp. 126-152.

Gårding, E. 1987. "Speech Act and Tonal Pattern in Standard Chinese: Constancy and Variation". *Phonetica,* 44, pp. 13-29.

Gass, S., and L. Selinker. (2008) *Second Language Acquisition: An Introductory Course*. 3rd Edition Hillsdale, Routledge (Taylor and Francis)

Gili Gaya, S. 1940. "La cantidad silábica en la frase". *Castilla (Valladolid)* 1, pp. 287-98 (Cited in Gutiérrez-Díez, 2001).

Goldsmith, J. 1976. "Autosegmental phonology", Doctoral dissertation, MIT, Cambridge, Mass.

Grabe, E. 2002. "Variation adds to prosodic typology". In *Proceedings of the Speech Prosody 2002 Conference,* 11-13 April 2002, edited by B. Bel and I. Marlin, pp. 127-132. Aix-en-Provence: Laboratoire Parole et Langage. http://www.phon.ox.ac.uk/~esther/Grabe.doc

Grabe, E., and E.L. Low. 2002. "Durational variability in speech and the rhythm class hypothesis". In *Papers in Laboratory Phonology 7*, edited by C. Gussenhoven and N. Warner. Berlin: Mouton de Gruyter. http://www.phon.ox.ac.uk/~esther/ivyweb/Grabe_Low.doc

Grabe, E., and B. Post. 2002. "Intonational variation in the British Isles". In Speech Perception-2002, pp. 343-346

Grabe, E., B. Poste, F. Nolan and K. Farrar. 2000. "Pitch accent realization in four varieties of British English". *Journal of Phonetics* 28, pp. 161-185.

Grand Dictionnaire Hachette Oxford français-anglais / anglais-français - CD-ROM PC. 2003. Oxford University Press.

Gu, H-Y., and K-H. Wan. 2004. "An acoustic and articulatory knowledge integrated method for improving synthetic mandarin speech's fluency". *International Symposium on Chinese Spoken Language Processing 2004, Hong Kong,* pp. 205-208.

Guilbault, Ch. 2002. "The acquisition of French rhythm by English second language learners". Doctoral Dissertation, University of Alberta.

Gut, U., J. Trouvain and W.J. Barry. 2007. "Bridging research on phonetic descriptions with knowledge from teaching practice -- The case of prosody in non-native speech". In: *Non-native prosody: phonetic description and teaching practice*, edited by J. Trouvain and U. Gut. Berlin: Mouton de Gruyter.

Gutiérrez-Diez, F. 2001. "The acquisition of English syllable timing by native Spanish speakers learners of English: an empirical study." *International Journal of English Studies (IJES)* 1:1, pp. 93-113.

Hayes, B.1980. *A metrical theory of stress rules.* Unpublished Doctoral Dissertation. MIT. (Cited in Pierrehumbert, 1980)

_____ . 1989. "The prosodic hierarchy in meter". In *Phonetics and Phonology, Vol 1: Rhythm and Meter*, edited by P. Kiparsky and G. Youmans, pp. 201-260. San Diego: Academic Press. (Cited in Shattuck-Hufnagel and Turk, 1996)

He, K., and D, Li. 1987. *Xiandai Hanyu san qian changyong ci biao [three thousand most commonly used words in modern Chinese].* Beijing: Beijing Shifan Daxue Chubanshe). (Cited in Duanmu, 2000).

Hirst, D.J. 2005. "Form and function in the representation of speech prosody". *Speech Communication,* 46, Issues 3-4, pp. 334-347.

Hockett, C. 1958. *A Course in Modern Linguistics*. New York: MacMillan.

Ho, A. 1977. "Intonation variation in a Mandarin sentence for three expressions: interrogative, exclamatory and declarative". *Phonetica*, 34, pp. 446-457.

Hu, M-Y. 1985. *Yuyan yu yuyanxue [Languages and Linguistics].* Wuhan: Hubei Jiaoyu Chubanshe. (Cited in Lee, 2005)

Hua, T-F. 2003. "The acquisition of English speech Rhythm by adult Chinese ESL and EFL learners". Doctoral thesis. Manoa: University of Hawaii.

Jarmulowicz, L. D. 2002. "English derivational suffix frequency and children's stress judgments". *Brain and Language, 81*, 192-204

Jin, S. 1996. "An acoustic study of sentence stress in Mandarin Chinese". Ph.D. dissertation, Ohio State University.

Jun, S-A. (ed.) 2005. *Prosodic Typology: The Phonology of Intonation and Phrasing*. New York: Oxford University Press.

Jun, S.-A., and C. Fougeron. 1995. "The accentual phrase and the prosodic structure of French". *Proceedings of International Congress of Phonetic Science 1995 conference (ICPhS 95)*, pp. 722-725.

_____ . 2000. "A phonological model of French intonation". In *Intonation: Analysis, Modeling and Technology*, edited by A. Botinis, pp. 209-242. Dordrecht: Kluwer Academic Publishers.

_____ . 2002. "Realizations of accentual phrase in French intonation". *Probus* 14, pp. 147-172.

Kaminskaïa, S. (forthcoming) "Étude comparée de l'intonation de la parole spontanée dans deux dialectes de français". Doctoral dissertation, Université Western Ontario.

_____ . 2005. "Une étude comparée d'un contour de continuité en français de France et du Québec". *Proceedings of the 2005 annual conference of the Canadian Linguistic Association.*

Kamiyama, T. 2004. "Perception of foreign accentedness in L2 prosody and segments: L1 Japanese speakers learning L2 French", In *Speech Prosody-2004*, pp. 721-724.

Keating, P., and C. Esposito. 2007. "Linguistic voice quality". *Department of Linguistics, UCLA. Working Papers in Phonetics*, Paper No105_6. pp. 85-91, http://repositories.cdlib.org/uclaling/wpp/No105_6

Keller, E. 2005. "A phonetician's view of signal generation for speech synthesis". *Studientexte zur Sprachkommunikation, 36,* pp. 13-20.

http://mypage.bluewin.ch/ekeller00/pdf.files/Keller_05_PhoneticiansView.pdf

Kochanski, G., C. Shih and H. Jing. 2003. "Hierarchical structure and word strength predication of Mandarin prosody", In *International J. Speech Technology* 6:1, pp. 33-43.

Krashen, S. 1981. *Second Language Acquisition and Second Language Learning.* Oxford: Pergamon.
http://www.sdkrashen.com/SL_Acquisition_and_Learning/index.html

Kunter, G., and I. Plag. 2007. "What is compound stress?", in *Proceedings of the 17th International Congress of Phonetic Sciences*, University of Saarbrücken, 6-10 August 2007. Saarbrücken: Universität Saarbrücken http://www2.uni-siegen.de/~engspra/DFG-Project/Plag-Kunter-2007.pdf

Labov, W., S. Ash and C. Boberg. 2006. *Atlas of North American English: Phonetics, Phonology and Sound Change.* Berlin/New York: Mouton de Gruyter.

Lacheret-Dujour, A., and F. Beaugendre. 1999. *La prosodie du français.* Paris: CNRS.

Ladd, D.R. 1983. Phonological features of intonational peaks. *Language,* 59, pp.721–759.

Ladefoged, P. 1975. *A Course in Phonetics.* Orlando: Harcourt Brace.

Lado, R. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor: University of Michigan Press.

Lane, L. 2005. *Focus on Pronunciation 2.* New York: Pearson.

Lane, L. 1993. *Focus on Pronunciation: Principles and Practice for Effective Communication.* New York: Longman.

Leben, W. 1971. "Suprasegmental and segmental representation of tone". *Studies in African Linguistics*, supp. 2, pp. 183-200.

Lee, O. J. 2005. "The prosody of questions in Beijing Mandarin", doctoral dissertation, Ohio State University.

Lehiste, I. 1977. "Isochrony reconsidered". *Journal of Phonetics, 5*, pp. 253-263.

Lin, H. 2001. *A Grammar of Mandarin Chinese.* Munich, Germany: Lincom Europa.

Low, E.L. 1998. "Prosodic prominence in Singapore English". Doctoral dissertation, University of Cambridge.

Low, E.L., and E. Grabe. 1995. "Prosodic patterns in Singapore English". *Proceedings of the International Congress of Phonetic Sciences, 3*, pp. 636-639.

Major, R. C. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology (Reflective Teaching and the Social Conditions of Schooling).* Mahwah, NJ: Lawrence Erlbaum Associates.

McCarthy 1982. "Nonlinear phonology: an overview". *GLOW Newsletter* 8, 63–77. http://www.meertens.knaw.nl/glow2002/mccarthy.pdf

Ménard, L., C. Ouellon and J. Dolbec. 1999. "Prosodic markers of regional group membership : the case of the French of Quebec versus France". *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, USA, pp.1601-1604.

Mertens, P. 1993. "Accentuation, intonation et morphosyntaxe". *Travaux de Linguistique* 26, pp. 21.69.

Miller, S. 2000. *Targeting Pronunciation: the Intonation, Sounds, and Rhythm of American English.* Boston, MA: Houghton Mifflin.

Miracle, W.C. 1989. "Tone production of American students of Chinese: A preliminary

acoustic study". *Journal of Chinese Language Teachers Association,* 24, 49-65. (in Wang, 2003)

Morton, J., S. Marcus and C. Frankish. 1976. "Theoretical Note, Perceptual Centers (Pcenters)". *Psychological Review,* 83, pp. 405-408. http://www.personal.rdg.ac.uk/~llsroach/phon2/pcent-xtrax.htm

Munro, M. J. 1995. "Nonsegmental factors in foreign accent: Ratings of filtered speech". *Studies in Second Language Acquisition,* 17, pp. 17–34.

Navarro Tomas, T. 1916. "Cantidad de las vocales acentuadas". In *Revista de Filología Española,* III, pp. 387-408. (Cited in Gutiérrez-Díez, 2001).

_____ . 1917. "Cantidad de las vocales inacentuadas". *Revista de Filología Española*, IV, pp. 371-388. (Cited in Gutiérrez-Díez, 2001).

_____ . 1921. "Historia de algunas opiniones sobre la cantidad silábica". *Revista de Filología Española*, VIII, pp. 30-52. (Cited in Gutiérrez-Díez, 2001).

_____ . 1922. "La cantidad silábica en unos versos de Rubén Darío". *Revista de Filología Española,* IX, pp. 1-29. (Cited in Gutiérrez-Díez, 2001).

Nespor, M., and I. Vogel. 1986. *Prosodic Phonology.* Dordrecht: Foris.

O'Connor, J. D., and G. F. Arnold. 1961. *Intonation of Colloquial English.* London: Longmans.

Otlowski, M. 1998. "Pronunciation: what are the expectations?". *The Internet TESL Journal*, Vol. IV:1, January 1998.

http://iteslj.org/Articles/Otlowski-Pronunciation.html

Pellegrino, F., J. Chaucat, R. Rakotomalala and J. Farinas. 2002. "Can automatically extracted rhythmic units discriminate among languages"?. *Proceedings of Prosody 2002,* Aix-en-Provence, France.

Peng, S-H., M. K. M. Chan, C-Y. Tseng, T. Huang, O-J. Lee and M. Beckman. 2005. "Towards a pan-Mandarin system for prosodic transcription". In *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by S-A. Jun, pp. 230-270. Oxford: Oxford University Press.

Pennington, M. 1989. "Teaching pronunciation from the top down". *Regional Language Centre (RELC) Journal*, 20:1, pp. 21-38.

Peperkamp, S., and E. Dupoux. 2002. "A typological study of stress 'deafness'". In *Laboratory Phonology 7,* edited by C. Gussenhoven and N. Warner, pp. 203-240. Berlin: Mouton de Gruyter.

Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation.* Doctoral Dissertation. MIT. Bloomington: Indiana University Linguistics Club.

Pike, K. 1945. *Intonation of American English.* University of Michigan Press: Ann Arbor.

Plag, I., and G. Kunter 2007. "The phonetics of primary vs. secondary stress in English". http://www.uni-siegen.de/~engspra/DFG-Project/Plag-Kunter-2007.pdf

Plag, I., G. Kunter, M. Braun and S. Lappe. (to appear). "Modeling compound stress in English".

Pointon, G. 1980. "Is Spanish really syllable-timed?". *Journal of Phonetics*, 8, pp. 293-304.

Port, R. F., K. Tajima and F. Cummins. 1998. "Speech and rhythmic behavior". In *Non-linear Developmental Processes,* edited by G. Savelsburgh, H. van der Maas and P. van Geert. pp. 53-78. Amsterdam: Elsevier.

Post, B. 2000. *Tonal and phrasal structures in French intonation.* Doctoral Dissertation, University of Nijmegen. The Hague: Holland Academic Graphics.

Prieto, P., J. van Santen and J. Hirschberg. 1995. "Tonal alignment patterns in Spanish". *Journal of Phonetics,* 23, pp. 429-451.

Ramus, F. 1997. "Le rôle du rythme pour la discrimination des langues". *Actes des JIOSC 97*, Orsay, 1-2/12/1997, pp. 225-229. http://www.ehess.fr/lscp/persons/ramus/docs/jiosc97.htm

_____ . 1999. "La discrimination des langues par la prosodie: Modélisation linguistique et études comportementales". In *De la caractérisation à l'identification des langues. Actes de la 1ère journée d'étude sur l'identification automatique des langues*, Lyon, 19 janvier 1999. Edited by F. Pellegrino, pp. 186-201. Lyon: Editions de l'Institut des Sciences de l'Homme

http://www.lscp.net/persons/ramus/docs/idlang99.pdf

_____ .. 2002. "Acoustic correlates of linguistic rhythm: perspectives". *Proceedings of the Speech Prosody 2002 Conference*, 11-13 April 2002, Aix-en-Provence: Laboratoire Parole et Langage, pp. 115-120. http://www.lscp.net/persons/ramus/docs/ramus_sp02.pdf

Ramus, F., M. D. Hauser, C. Miller, D. Morris and J. Mehler. 2000. "Language discrimination by human newborns and by cotton-top tamarin monkeys". *Science,* 288, pp. 349-351. http://cogprints.org/870/0/tamarin99.pdf

Roach, P. 1982. "On the distinction between 'stress-timed' and 'syllable-timed' languages". In *Linguistic controversies,* edited by D. Crystal, pp. 73-79. London: Edward Arnold.

Rossi, M. 1985. "L'intonation et l'organisation de l'énoncé". *Phonetica*, 42:2-3, pp. 135-153. (Cited in Lacheret-Dujour, A., and F. Beaugendre. 1999).

Santerre, L. 1990. "La condition de non-contiguïté accentuelle en français: théorie et pratique". *Revue québécoise de linguistique*, 19:2, pp. 39-58.

Saville-Troike, M. 2006. *Introducing Second Language Acquisition.* Cambridge: Cambridge University Press.

Selinker, L. 1972. "Interlanguage". *International Review of Applied Linguistics in Language Teaching (IRAL)* 10, pp. 209-231.

Selkirk, E. 1980. "The role of prosodic categories in English word stress". *Linguistics Inquiry*, 11:3, pp. 563-606. (Cited in Pierrehumbert, 1980)

_____. 1982. *The Syntax of Words.* Cambridge, MA. MIT Press.

_____ . 1986. "On derived domains in sentence phonology". *Phonology Yearbook,* 3, pp. 371– 405.

Shattuck-Hufnagel, S., and A. Turk 1996. "A prosody tutorial for investigators of auditory sentence processing". *Psycholinguistic Research*, 25, pp. 193-247.

Shen, X. 1990. *The Prosody of Mandarin Chinese.* Berkeley: University of California Press.

Shen, J. 1985. "Beijing hua shengdiao de yinyu he yudiao [Pitch Range and Intonation of the Beijing Mandarin]". In: *Beijing yuyin shiyan lu.* Edited by T. Lin and L. Wang. Beijing: Peking University. (Cited in Lee, 2005)

Shih, C. 1997. "Mandarin third tone sandhi and prosodic structure". In *Studies in Chinese Phonology*, edited by J. Wang and N. Smith, pp. 81-123. Berlin and New York: Mouton de Gruyter.

_____ . 2000. "A declination model of Mandarin Chinese". In *Intonation: Analysis, Modelling, and Technology.* Edited by A. Botinis, pp. 243-268. Kluwer Academic Publishers.

Silverman, K., and J. Pierrehumbert. 1990. "The timing of prenuclear high accents in English". *Papers in Laboratory Phonology I*, pp. 72-106. Cambridge: Cambridge University Press.

K. Silverman, M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert and J. Hirschberg. 1992. "ToBI: a standard for labeling English prosody". *Proceedings of the 1992 International Conference on Spoken Language Processing, ICSLP*, 2, pp. 867-870, 12-16 October 1992, Banff, Canada.

Stern, H.H. 1992. *Issues and Options in Language Teaching.* Oxford: Oxford University Press.

Suter, R., and E. Purcell. 1980. "Predictors of pronunciation accuracy: A re-examination". *Language Learning*, 30:2, pp.271-287.

Tajima, K. 1998. "Speech rhythm in English and Japanese: experiments in speech cycling". Doctoral dissertation, Bloomington: Indiana University.

Tajima, K., and R. F. Port. 2003. "Speech rhythm in English and Japanese". In *Phonetic Interpretation: Papers in Laboratory Phonology VI,* edited by J. Local, R. Ogden and R. Temple, pp.317-334. Cambridge: Cambridge University Press.

Trofimovich, P and W. Baker 2006. "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech". Studies *in Second Language Acquisition,* 28:1, pp. 1-30. http://education.concordia.ca/~pavel/pdf/Trofimovich-Baker-2006.pdf

Trouvain, J., and U. Gut. (eds.) 2007. *Non-native Prosody: Phonetic Description and Teaching Practice.* Berlin: Mouton de Gruyter.

Trudgill, P., and J. Hannah. 2002. *International English: A Guide to the Varieties of Standard English*, 4th edition. London: Arnold.

Tseng, C. 2002. "The prosodic status of breaks in running speech: examination and evaluation". *Speech Prosody 2002*, Aix-en-Provence, France, pp. 667-670.

_____. 2003. "Towards the role of Mandarin speech prosody: units, boundaries and their characteristics", *XIV Proceedings of International Conference of Phonetics Sciences,* Aug.1-9. Barcelona, Spain, pp. 481-484.

http://www.ntnu.edu.tw/tcsl/Seminar%20in%20Chinese%20Phonetics/0531/Required%20reading/1.pdf

Tseng, C., and Y-L. Lee. 2004. "Speech rate and prosody units: evidence of interaction from Mandarin Chinese". *Speech Prosody 2004*, Nara, Japan.

Ueyama, M. 2000. "Prosodic Transfer: An Acoustic Study of L2 English vs. L2 Japanese". Doctoral thesis. UCLA.

Vaissière, J. 1991. "Rhythm, accentuation and final lengthening in French". In *Music, Language, Speech and Brain.* J. Sundberg, L. Nord and R. Carlson, Macmillan Press, pp. 108-120.

_____. 2002. "Cross-linguistic prosodic transcription". In *Problems and methods in experimental phonetics* (Problem I metody eksperimental'no-fonetcheskih issledovanij), edited by N.B. Volskaja, P.A. Skrelin, N.D. Svetozarova, in honour of the 70th anniversary of Prof. L.V. Bondarko. St.-Perersburg, St.-Petersburg State University. pp.147-164.

http://www.personnels.univ-paris3.fr/users/vaissier/pub/ARTICLES/index_fichiers/2002.pdf

Vaissière, J., and A. Michaud. 2006. "Prosodic constituents in French: a data-driven approach". In *Prosody and Syntax*, edited by I.Fónagy, Y. Kawaguchi and T. Moriguchi. "Usage-based linguistic informatics" series, Amsterdam:John Benjamins.

http://www.cavi.univ-paris3.fr/ilpga/ED/student/stam/prosodic-constituents.pdf

Vennemann, T. 1988. *Preference Laws for Syllable Structure.* Berlin: Mouton-de Gruyter. (Cited in Auer, 1993)

Wang, W. S.Y. 1967. "Phonological features of tone". *International Journal of American Linguistics,* 33:2, pp. 93-105.

Wang, Y., J. A. Sereno and A. Jongman. 2006. "L2 acquisition and processing of Mandarin tones". In *The Handbook of East Asian Psycholinguistics: 1, Chinese* edited by P. Li, L-H. Tan, E. Bates, O. J. L. Tzeng. Cambridge: Cambridge University Press.

Wang, Y., A. Jongman and J. Sereno, 2003. "Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training". *Journal of the Acoustical Society of America*, 113, pp. 1033-1043.

Welby, P. 2004. "The structure of French intonational rises: A study of text-to-tune alignment". In *Proceedings of the Speech Prosody 2004 Conference,* edited by B. Bel and I. Marlien, pp. 127 – 130. Nara, Japan.

_____. 2006. "French intonational structure: evidence from tonal alignment". *Journal of Phonetics* 34:3, pp. 343–371.

http://www.icp.inpg.fr/%7Ewelby/PAPERS/WelbyJPhon2006.pdf

Wells, J.C. 2000. *Longman Pronunciation Dictionary.* Second edition. Harlow: Pearson Education Limited.

Lacheret-Dujour, A., and F. Beaugendre. 1999. *La prosodie du français*, Paris: CNRS Edition.

White, L. 1989. *Universal Grammar and Second Language Acquisition.* Amsterdam/Philadelphia: John Benjamins.

Whiteside, S. P. 1998. "Simulated Emotions: An Acoustic Study of Voice and Perturbation Measures". *Proceedings of the International Conference on Spoken Language Processing (ICSLP) 1998*, pp. 699-703.

Williams, E. 1976. 'Underlying tone in Margi and Igbo'. *Linguistic Inquiry*, 7, pp. 463-484.

Woo, N. 1969. "Prosody and phonology". Doctoral dissertation. Massachusetts Institute of Technology.

Wu, Z-J. 1982. "Putonghua yuju zhong de shengdiao bianhua [Tonal Changes in Mandarin Sentences]". *Zhongguo yuwen* 1982.6, pp. 439-450.

Xu, C. X., and Y. Xu, 2003. "Effects of Consonant Aspiration on Mandarin Tones". *Journal of the International Phonetic Association,* 33, pp. 165-181.

Xu, Y. 2001. "Fundamental frequency peak delay in Mandarin". *Phonetica,* 58, pp. 26-52. http://pc79-61.phon.ucl.ac.uk/yispapers/Peak-delay.pdf

_____. 2004. "Separation of functional components of tone and intonation from observed F0 Patterns". In: *Traditional Phonology to Modern Speech Processing: Festschrift for Professor Wu Zongji's 95th Birthday*, edited by G. Fant, H. Fujisaki, J. Cao, and Y. Xu, pp. 485-505. Beijing Foreign Language Teaching and Research Press.
http://pc79-61.phon.ucl.ac.uk/yispapers/Xu_Festschrift_for_Wu.pdf

_____. 2005. "Speech melody as articulatorily implemented communicative functions". *Speech Communication,* 46, pp. 220-251.
http://pc79-61.phon.ucl.ac.uk/yispapers/Xu_SpCm2005.pdf

Xu, Y. and F. Liu. 2006. "Tonal alignment, syllable structure and coarticulation: toward an integrated model". *Italian Journal of Linguistics* 18, pp. 125-159.
http://pc79-61.phon.ucl.ac.uk/yispapers/Xu_Liu_IJL2006.pdf

Xu, Y. and C. X. Xu. 2005. "Phonetic realization of focus in English declarative intonation". *Journal of Phonetics* 33: 159-197 http://pc79-61.phon.ucl.ac.uk/yispapers/Xu_Xu_JPhon2005.pdf

Yang, Y. and B. Wang. 2002. "Acoustic correlates of hierarchical prosodic boundary in Mandarin". In *Speech Prosody 2002*, pp. 707-710.

Yip, M. 2002. *Tone*. Cambridge, UK: Cambridge University Press.

Yuan, J. 2004. "Intonation in Mandarin Chinese: acoustics, perception, and computational modeling". Doctoral dissertation. Ithaca, New York, Cornell University.

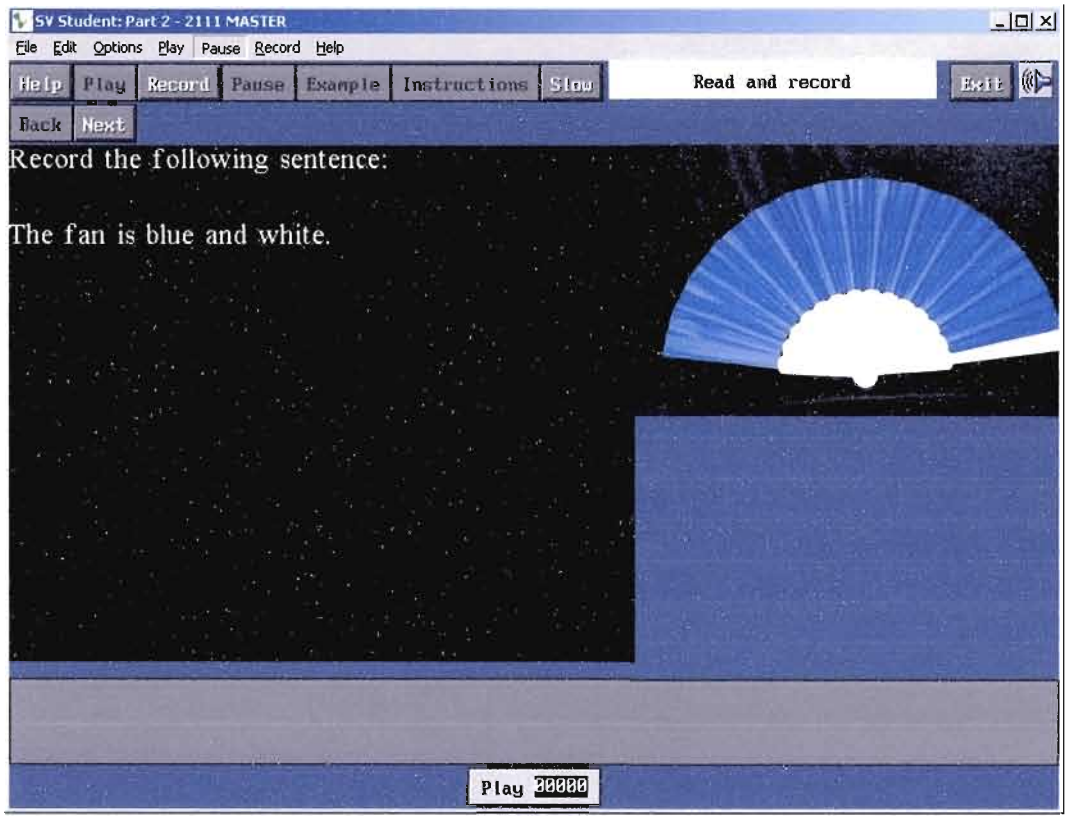
Yuan, J., L. Shen and F. Chen. 2002. "The acoustic realization of anger, fear, joy and sadness in Chinese". In *Proceedings of Seventh International Conference on Spoken Language Processing*, Denver, Colorado, pp. 2025–2028.

Zawaydeh B. A., Tajima, K. and M. Kitahara. 2002. "Discovering Arabic Rhythm through a Speech Cycling Task". In *Perspectives on Arabic Linguistics XIII-XIV,* edited by D.B. Parkinson and E. Benmamoun, pp. 39-58. Amsterdam/ Philadelphia: John Benjamins, http://www.f.waseda.jp/kitahara/Paper/ZTK02.pdf

Zeng, X.L., P. Martin and G. Boulakia. 2004. "Tones and intonation in declarative and interrogative sentences in Mandarin". In *Tonal Aspects of Languages (TAL) 2004*, pp. 235-238.

APPENDIX 1

Part 2 of the test: subjects were asked to produce the sentence that would appear on the screen. A picture illustrating the sentence was provided to ensure the participants understood the sentences. The recordings were done on Can8.
The elicitation sentences and pictures are provided below.



The fan is blue and white.

The girl holds a fan.

The first fan on the left is red.

There are three fans: a black fan, a white fan, and a red fan.

The old man holds a fan in his right hand.

The girl holds a pink fan.

The fan club cheered when the team won.

He received a large bag of fan mail.

The black cat plays in the snow.

The cat sleeps on the roof of the car.

The children made a snow cat.

The cats sleep on the roof.

The cattleman rides across the plains.

Catchers wear special equipment to help prevent injury.

Part 4 of the test: subjects were asked to produced the sentence that would appear on the screen. No picture was provided. The recordings were done on Can8.
The elicitation sentences are provided below.

Prison offered him the perfect opportunity to study.

The carpenter is impatient and motivated.

They evaluate the politician's attitude.

Appendix 2

Elicitation sentences in Standard Chinese. The presentation of the sentences was as follows:

The entire group of elicitation sentences is provided below with the Pinyin version. The **titles** and highlights of "fan" are provided here to facilitate understanding but were not provided to subjects so as to not draw their attention to the words.

Fàn (tone 4)

First syllable of the sentence

饭吃完了，菜还有一些。
(fàn chī wán le, cài hái yǒu yì xiē.)
When the rice is eaten, the rest of the meal is not.

Mid sentence/phrase:

他喝完酒以后，没吃饭就走了
(tā hē wán jiǔ yǐ hòu, méi chī fàn jiù zǒu le.)
He left after drinking alcohol and without eating.

Last syllable of an affirmative sentence:

他光吃菜，不吃饭
(tā guāng chī cài, bù chī fàn.)
He only eats vegetables, not rice.

Before a comma : 我们 去 吃饭 ，好不好？
(wǒ men qǔ chī fàn, hǎo bù hǎo)
Let's go for a meal, shall we?

Last syllable of an interrogative sentence without the syntactic marker "ma".
你昨天晚上吃的什么饭？
(nǐ zuó tiān wǎn shàng chī de shén me fàn?)
What did you have to eat last night?

fān (tone 1)

First syllable of the sentence: 帆 船 运 动 很 有 趣
fān chuán yùn dòng hěn yǒu qù.
Yachting is very interesting.

Mid sentence/phrase: 我 喜 欢 帆 船 运 动
wǒ xǐ huān fān chuán yùn dòng
I like yachting.

Last syllable of an affirmative sentence:
一 定 要 撑 住 帆
yí dìng yào chēng zhù fān
Someone needs to hold the sail.

Before a comma : 不 但 要 撑 住 帆 , 还 要 把 好 舵。
Bù dān yào chēng zhù fān, hái yào bǎ hǎo duò.
Not just to hold the sail, but also manage the tiller.

Last syllable of an interrogative sentence without the syntactic marker "ma".
你 能 不 能 撑 帆 ？
nǐ néng bù néng chēng fān?
Can you hold the sail?

fán (tone 2)

First syllable of the sentence: 繁 琐 的 手 续 让 他 头 疼
Fán suǒ de shǒu xù ràng tā tóu téng
The tedious procedure gave him a headache

Mid sentence/phrase:: 登 记 手 续 的 繁 琐 让 他 头 疼。
Dēng jì shǒu xù de fán suǒ ràng tā tóu téng.
The complication of the registration procedure gave him a headache.

Last syllable of an affirmative sentence:

公 司 的 事 让 他 心 烦
gōngsī de shì ràng tā xīn fán
Company business annoyed him.

Before a comma :

不 能 忍 受 平 凡 , 就 难 以 成 功
bù néng rěn shòu píng fán, jiu nán yǐ chéng gōng
It is difficult to succeed if you cannot endure the ordinary.

Last syllable of an interrogative sentence without the syntactic marker "ma".

他 是 不 是 很 平 凡
tā shì bú shì hěn píng fán?
Isn't he very ordinary?

fǎn (tone 3)

First syllable of the sentence:

返 程 的 票 也 订 好 了
Fǎn chéng de piào yě dìng hǎo le
The return ticket was also booked.

Mid sentence/phrase:

他 没 有 预 定 返 程 票
tā méi yǒu yù dìng fǎn chéng piào
He has not booked his return ticket.

Last syllable of an affirmative sentence:

这 个 价 格 包 括 往 返
zhè gè jià gé bāo kuò wǎng fǎn
This price is for a round trip.

Before a comma :

如 果 包 括 往 返 , 这 个 价 格 就 不 算 贵
rǔ guǒ bāo kuò wǎng fǎn, jiu bú suàn guì.
If it is for a round-trip, the price is not too high.

Last syllable of an interrogative sentence without the syntactic marker "ma".

他 们 是 不 是 要 谋 反
tā mén shì bú shì yào móu fǎn?
They must plot a coup?

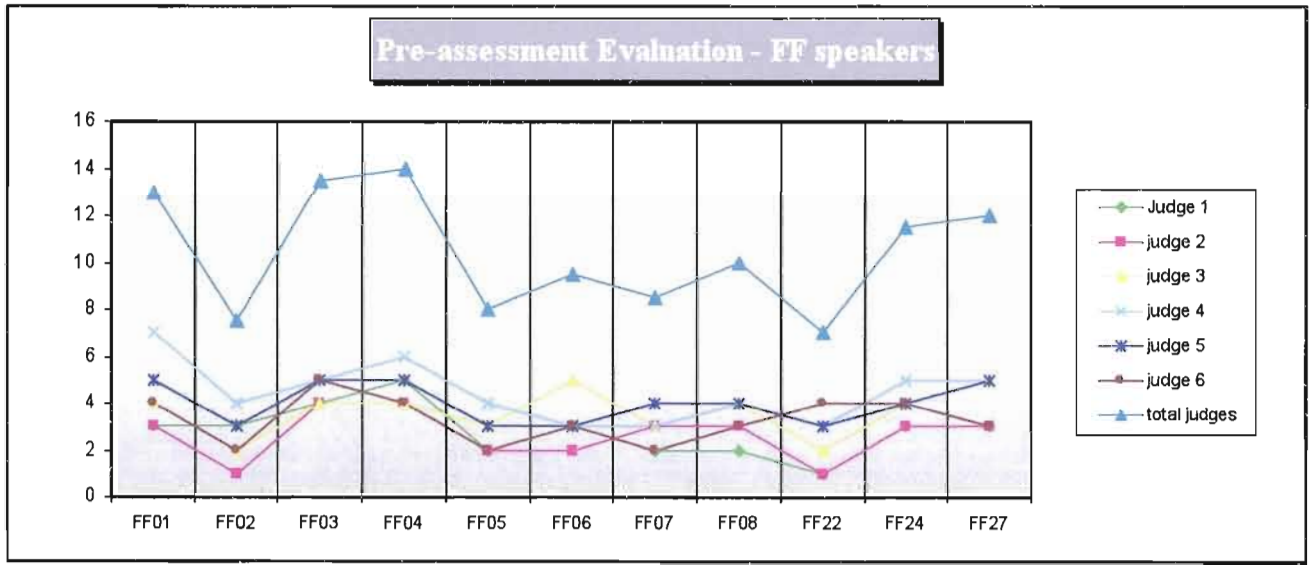
APPENDIX 3

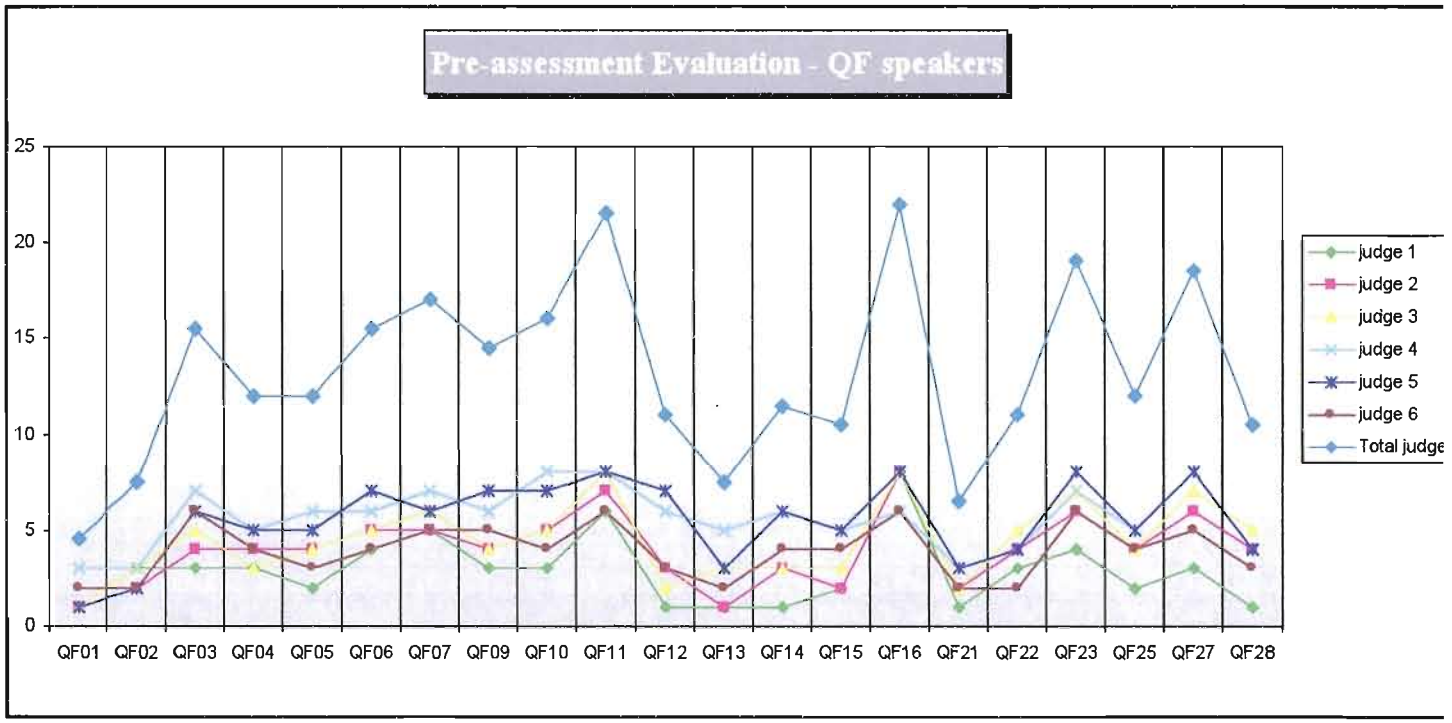
The following three graphs represent the assessment of each judge for each speaker of each linguistic group. The judges' choice of level was converted into points:

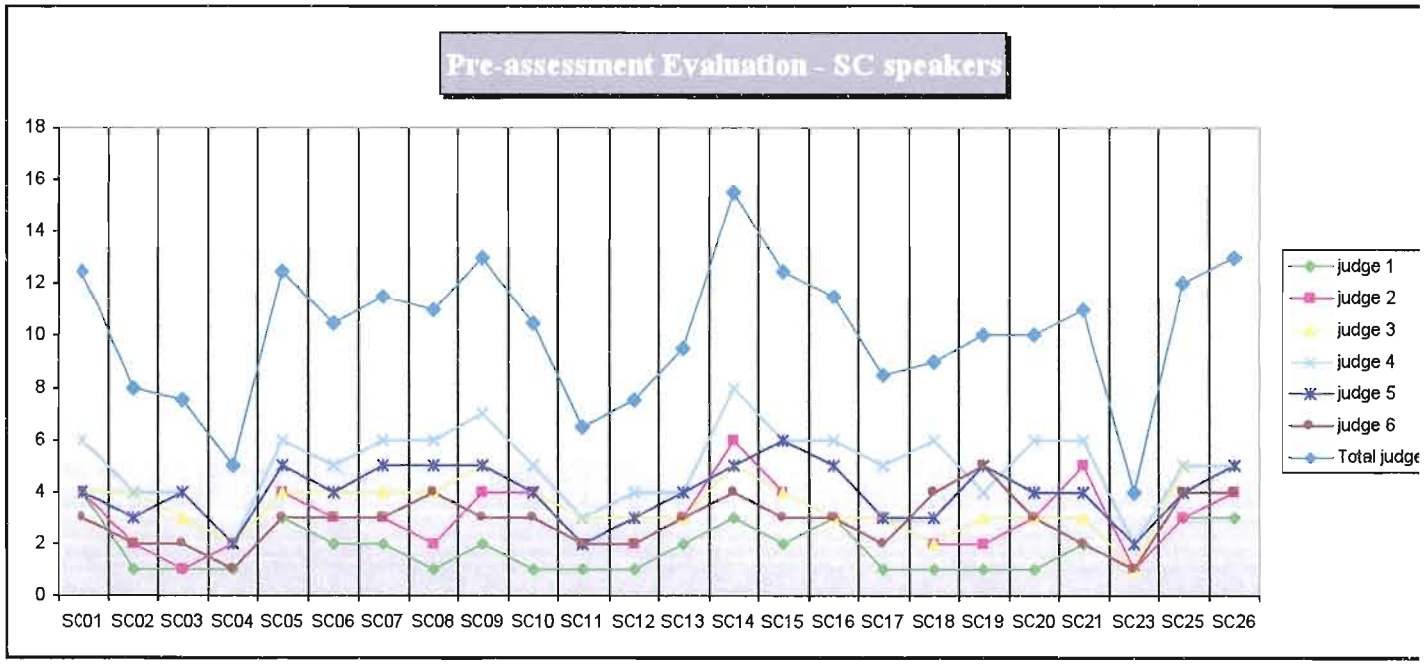
1) Beginner 1 point
2) Elementary (low) 2 points
3) Elementary (mid) 3 points
4) Elementary/intermediate 4 points
5) Intermediate (mid) 5 points
6) Intermediate (high) 6 points
7) Advanced 7 points
8) High advanced/fluent 8 points

The "total judges" represents the sum of all these points. For practical reasons of presentation, this number was divided by 2.

"Total judges" = [judge 1 + judge 2 + judge 3 + judge 4 + judge 5 + judge 6] / 2



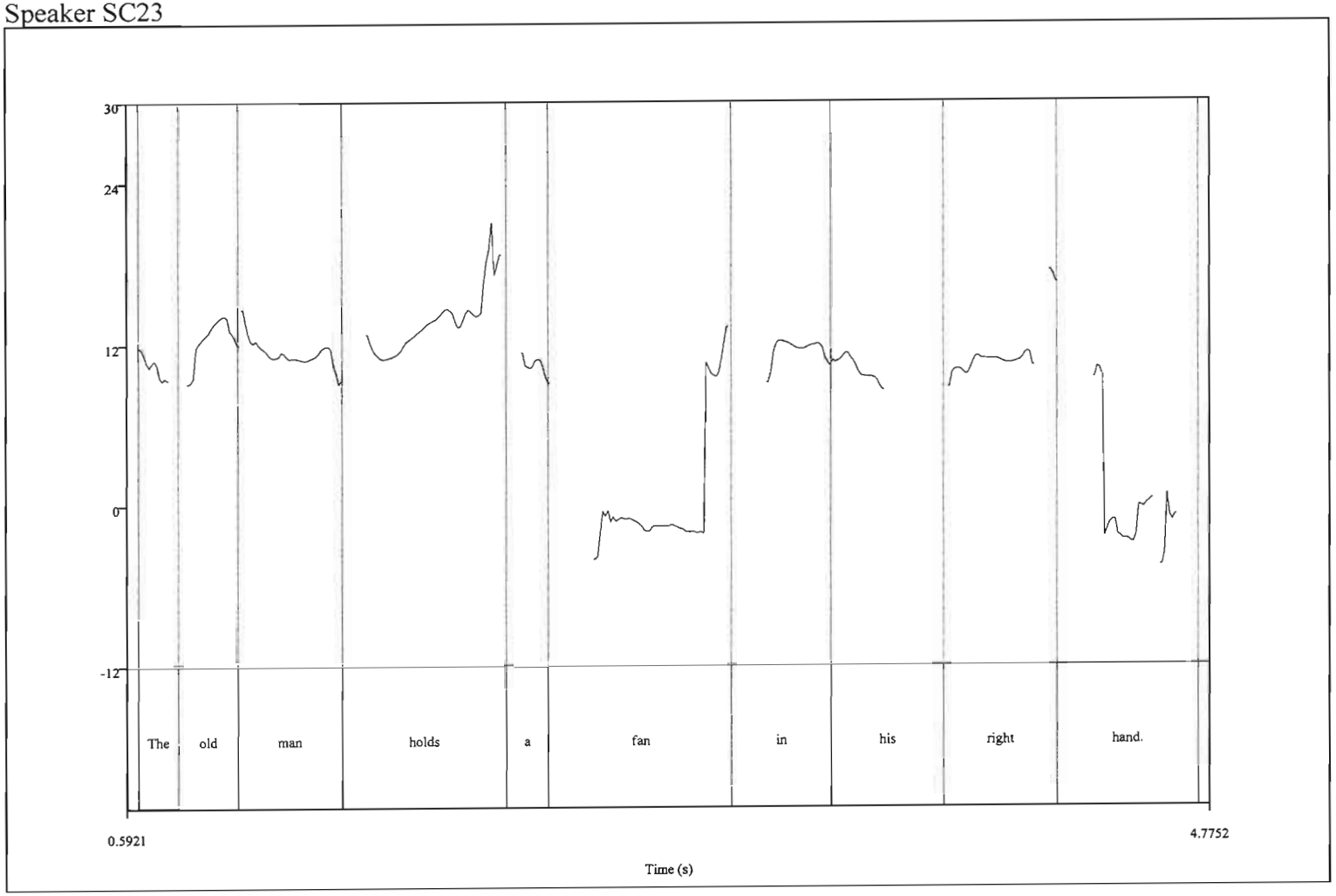

Pre-assessment Evaluation - QF speakers
25
20
15
10
5
0
QF01 QF02 QF03 QF04 QF05 QF06 QF07 QF09 QF10 QF11 QF12 QF13 QF14 QF15 QF16 QF21 QF22 QF23 QF25 QF27 QF28
judge 1
judge 2
judge 3
judge 4
judge 5
judge 6
Total judge


Pre-assessment Evaluation - SC speakers
18
16
14
12
10
8
6
4
2
0
SC01 SC02 SC03 SC04 SC05 SC06 SC07 SC08 SC09 SC10 SC11 SC12 SC13 SC14 SC15 SC16 SC17 SC18 SC19 SC20 SC21 SC23 SC25 SC26
judge 1
judge 2
judge 3
judge 4
judge 5
judge 6
Total judge

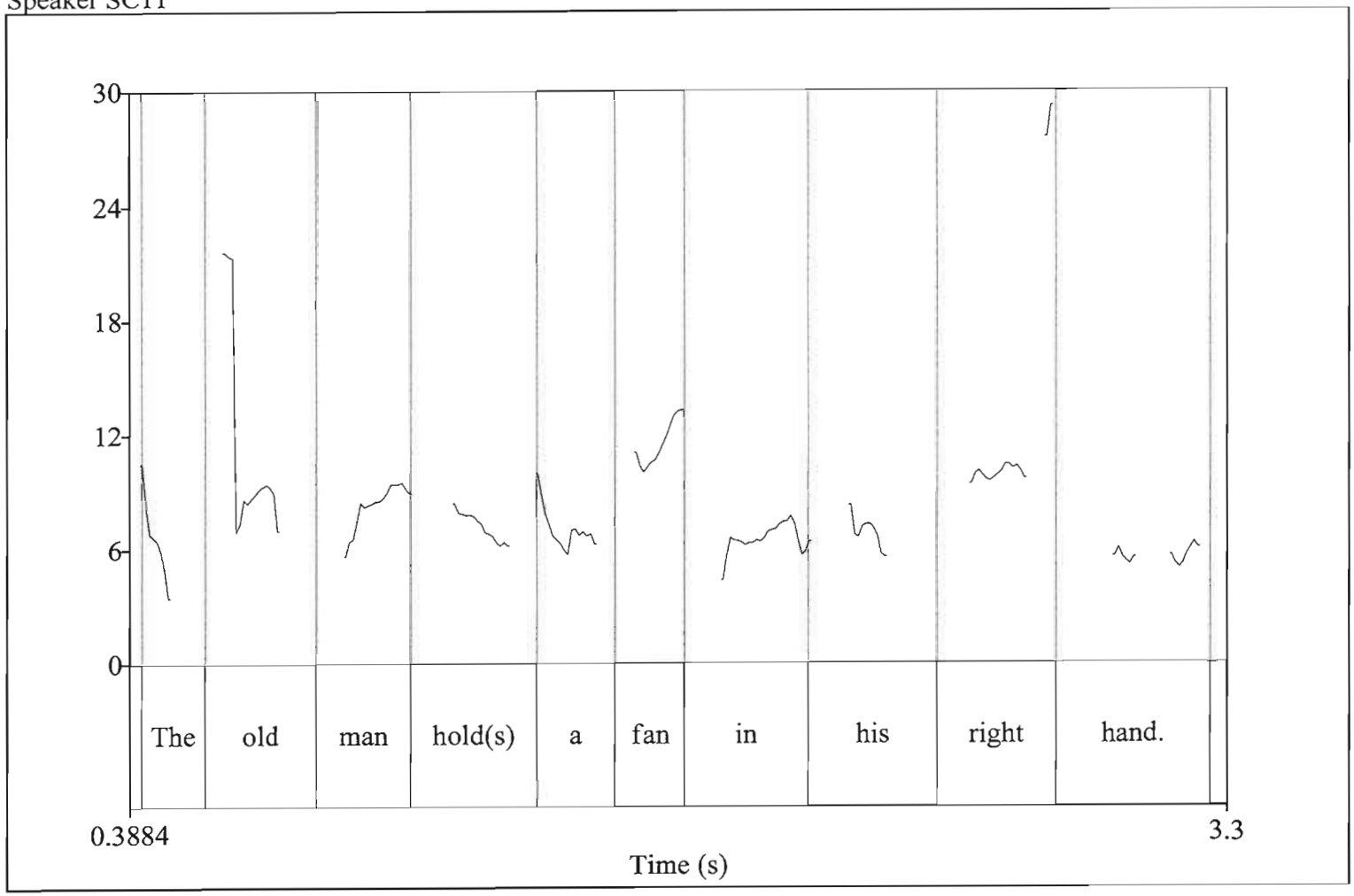
APPENDIX 4

Large-scale of F0 contours for experiment 4.
Subjects organized in ranking order.

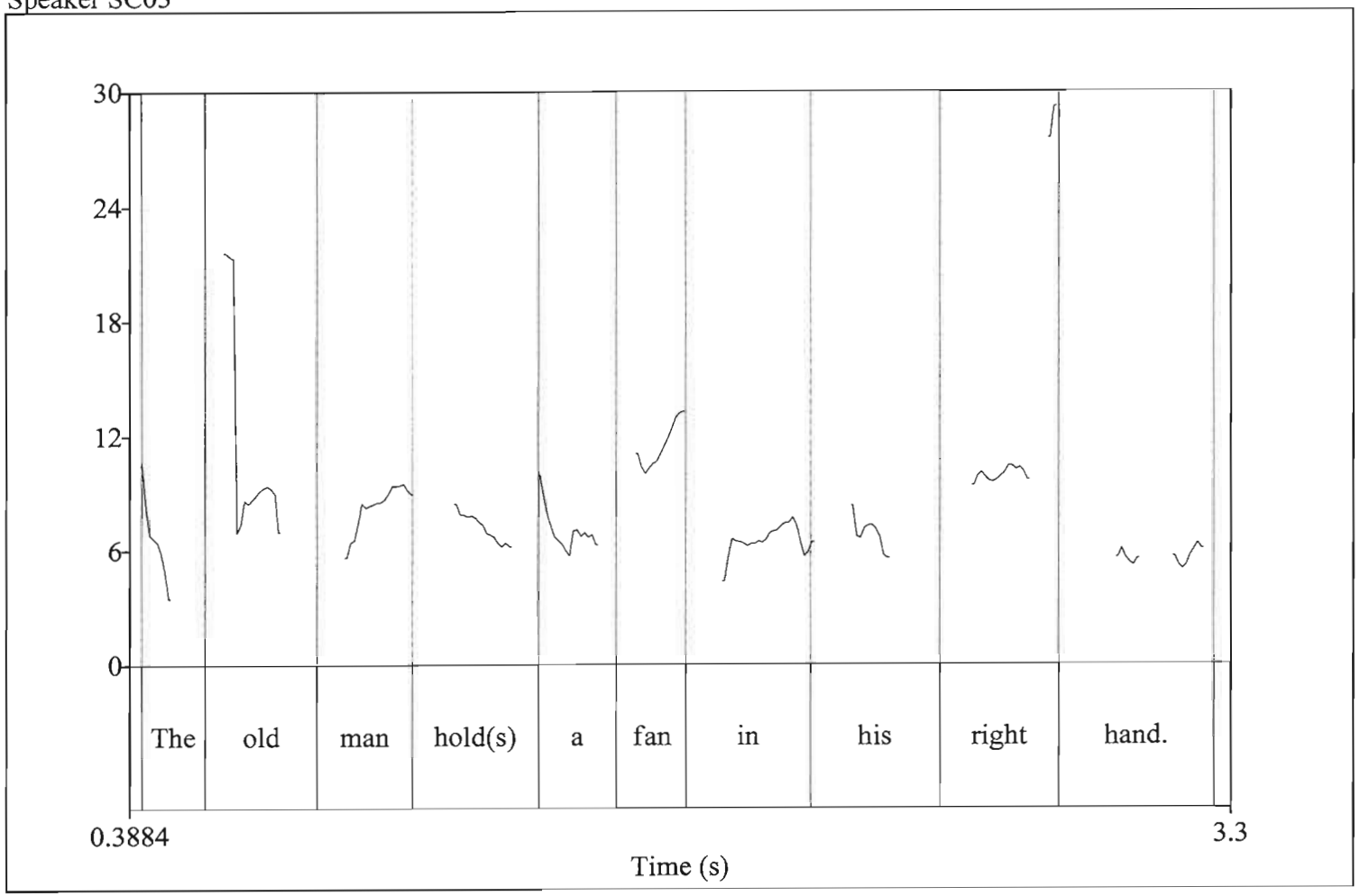
Speaker SC23



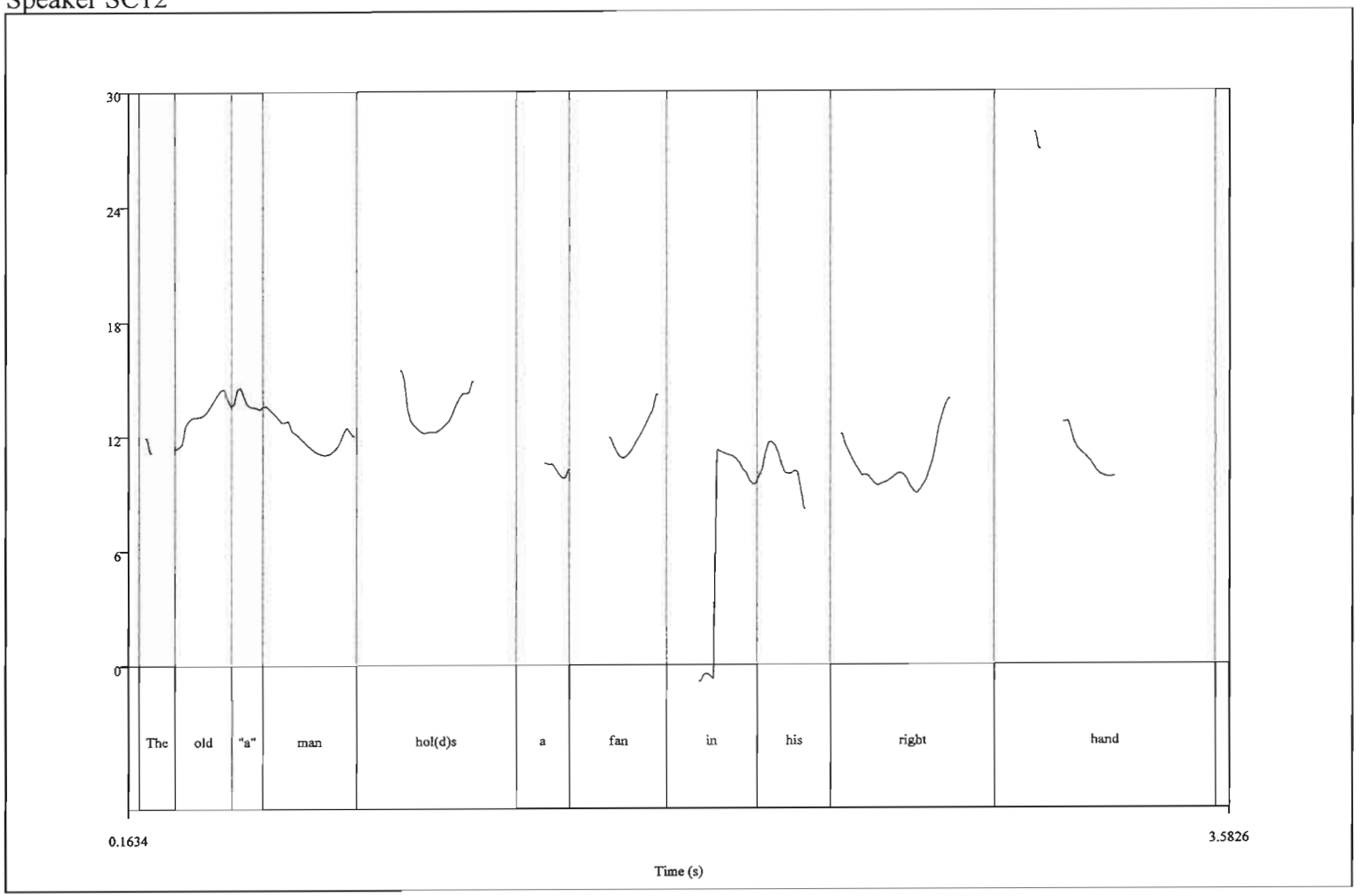
Speaker SC11



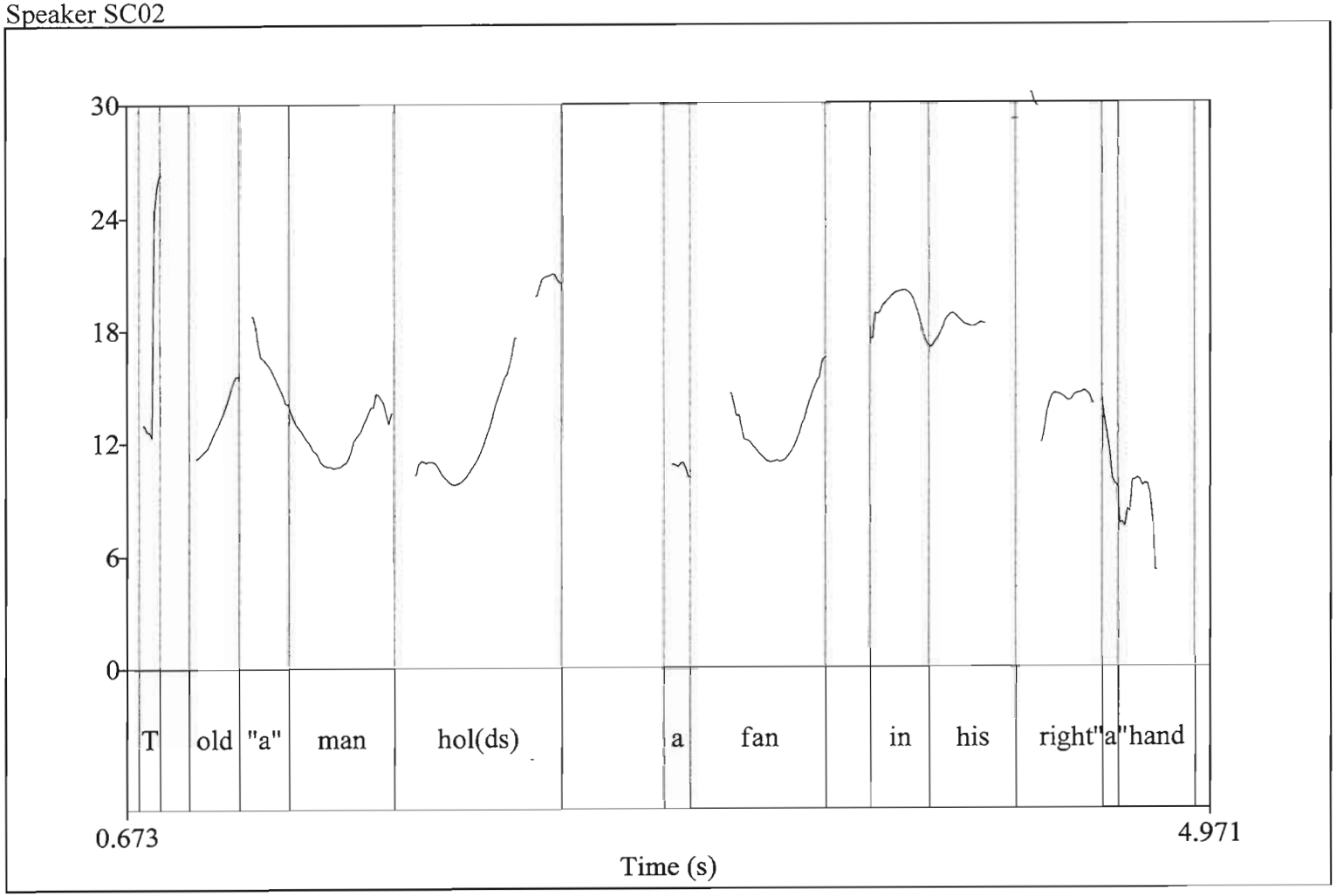
Speaker SC03



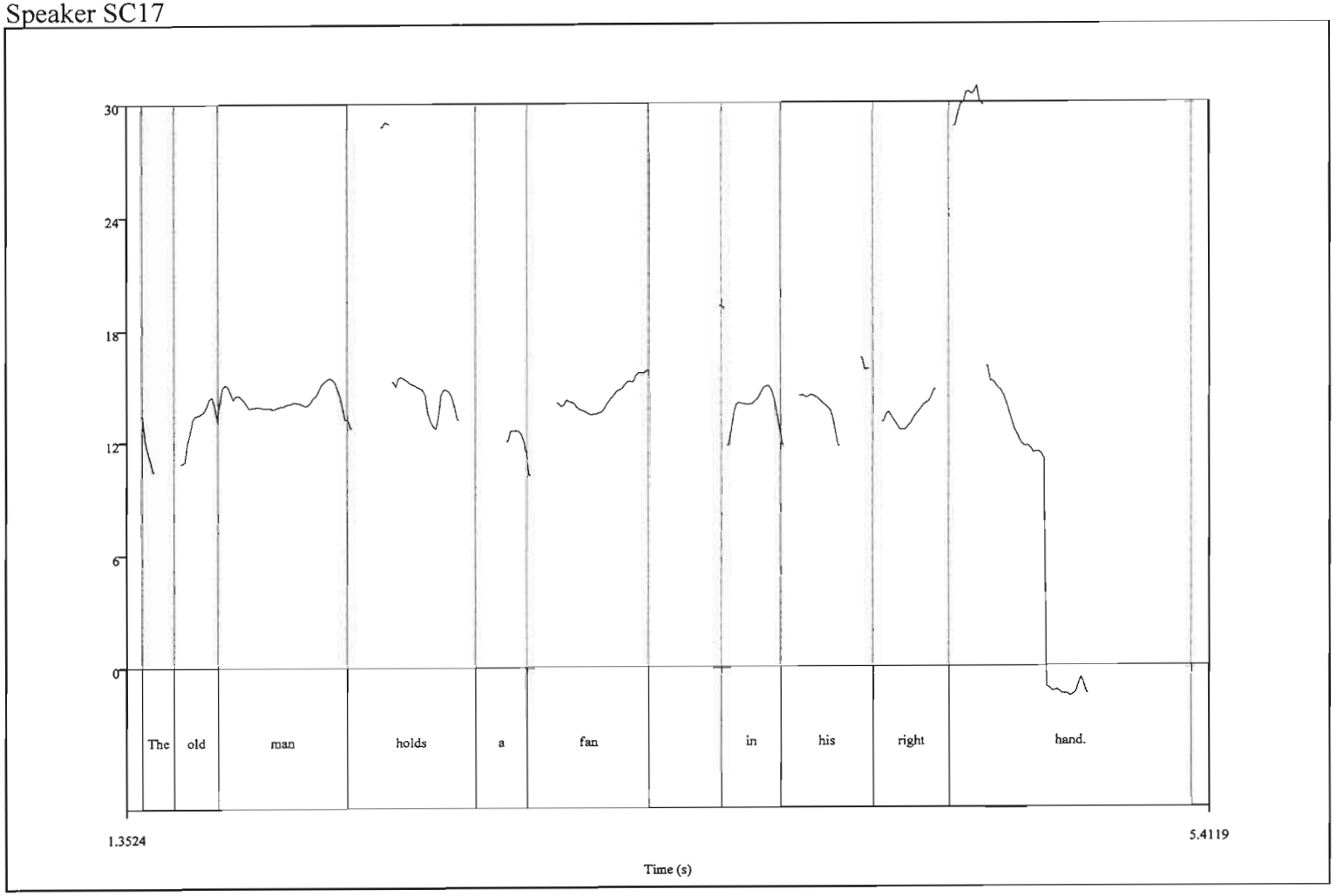
Speaker SC12



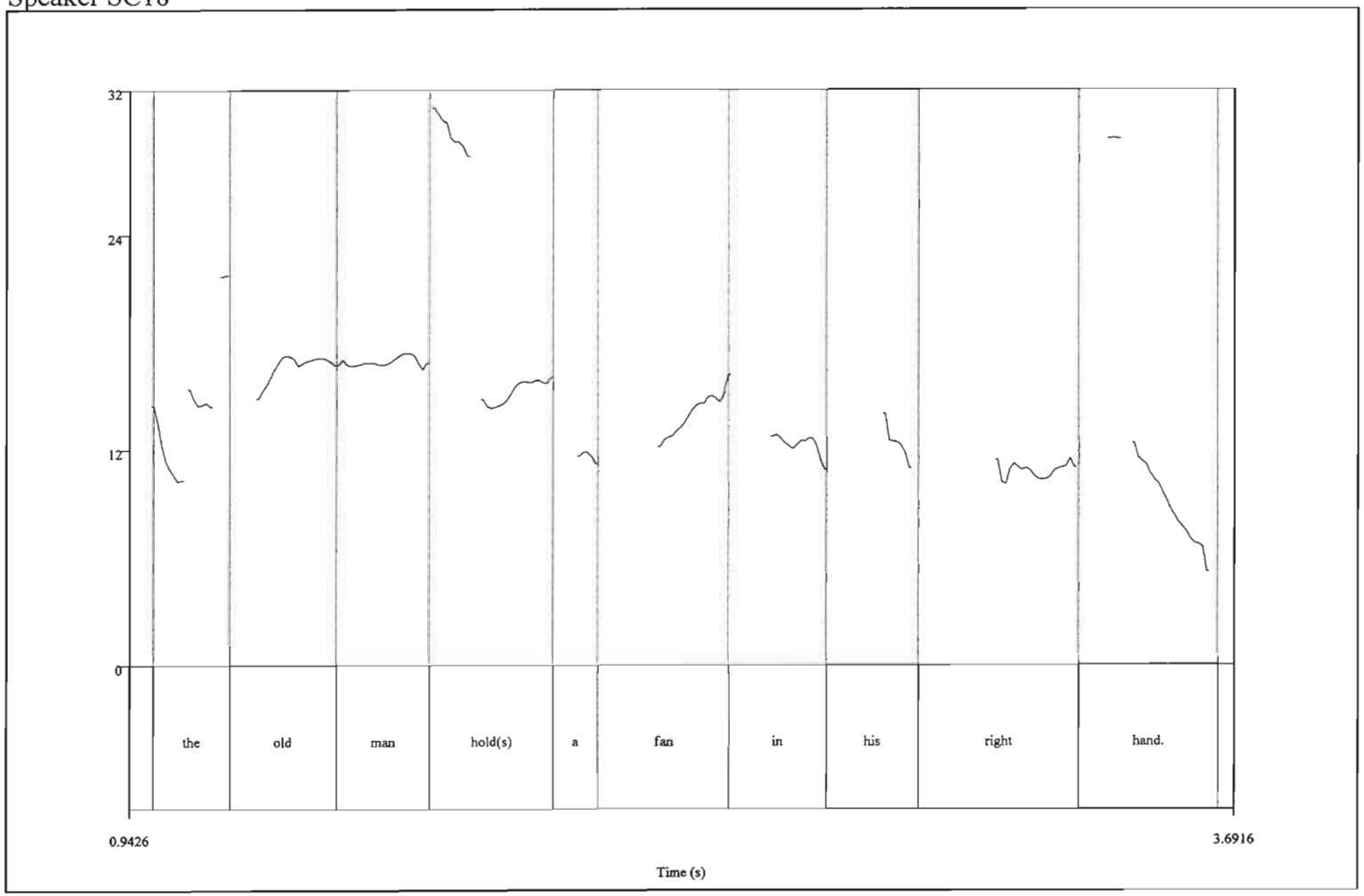
Speaker SC02



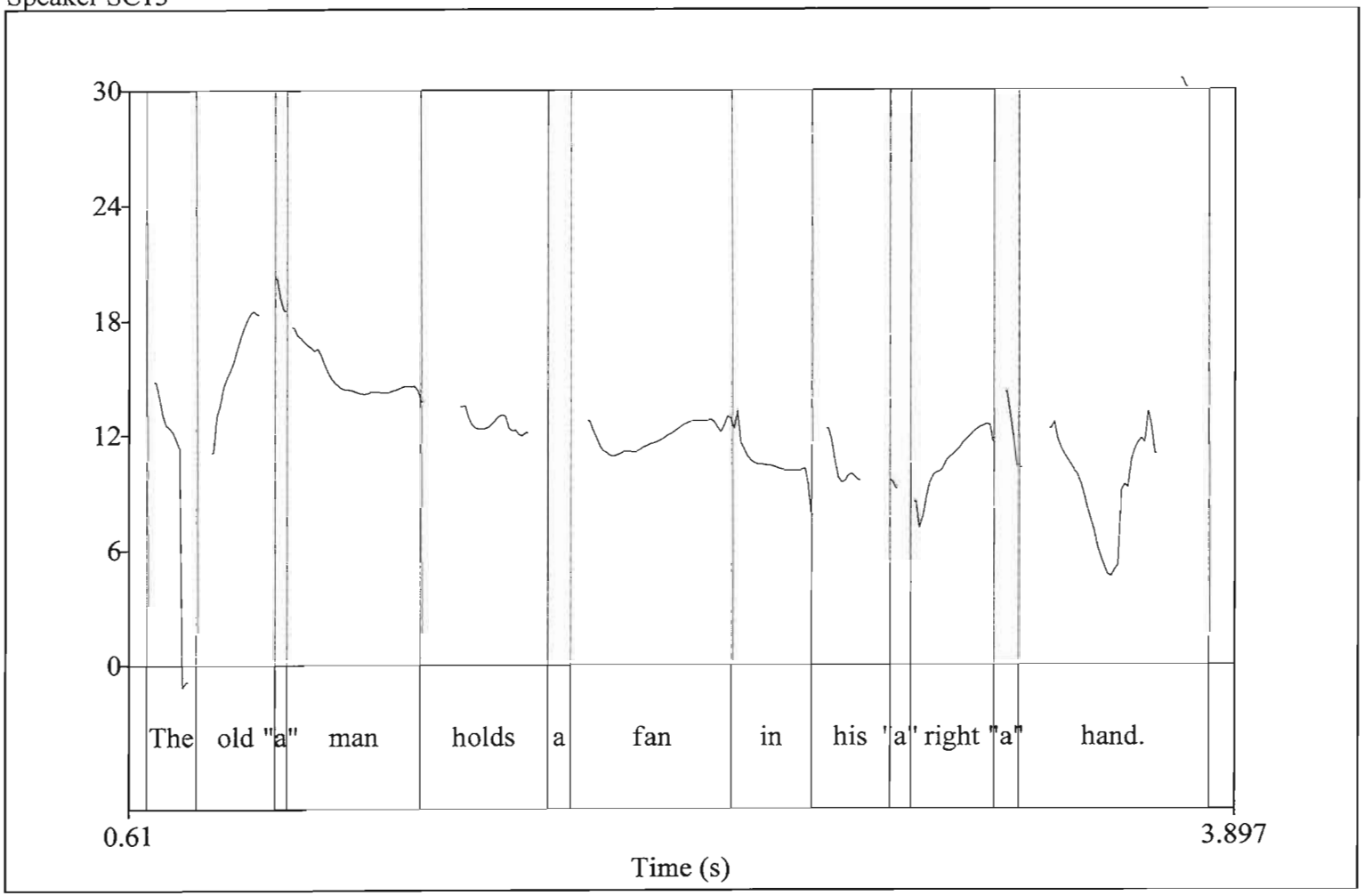
Speaker SC17



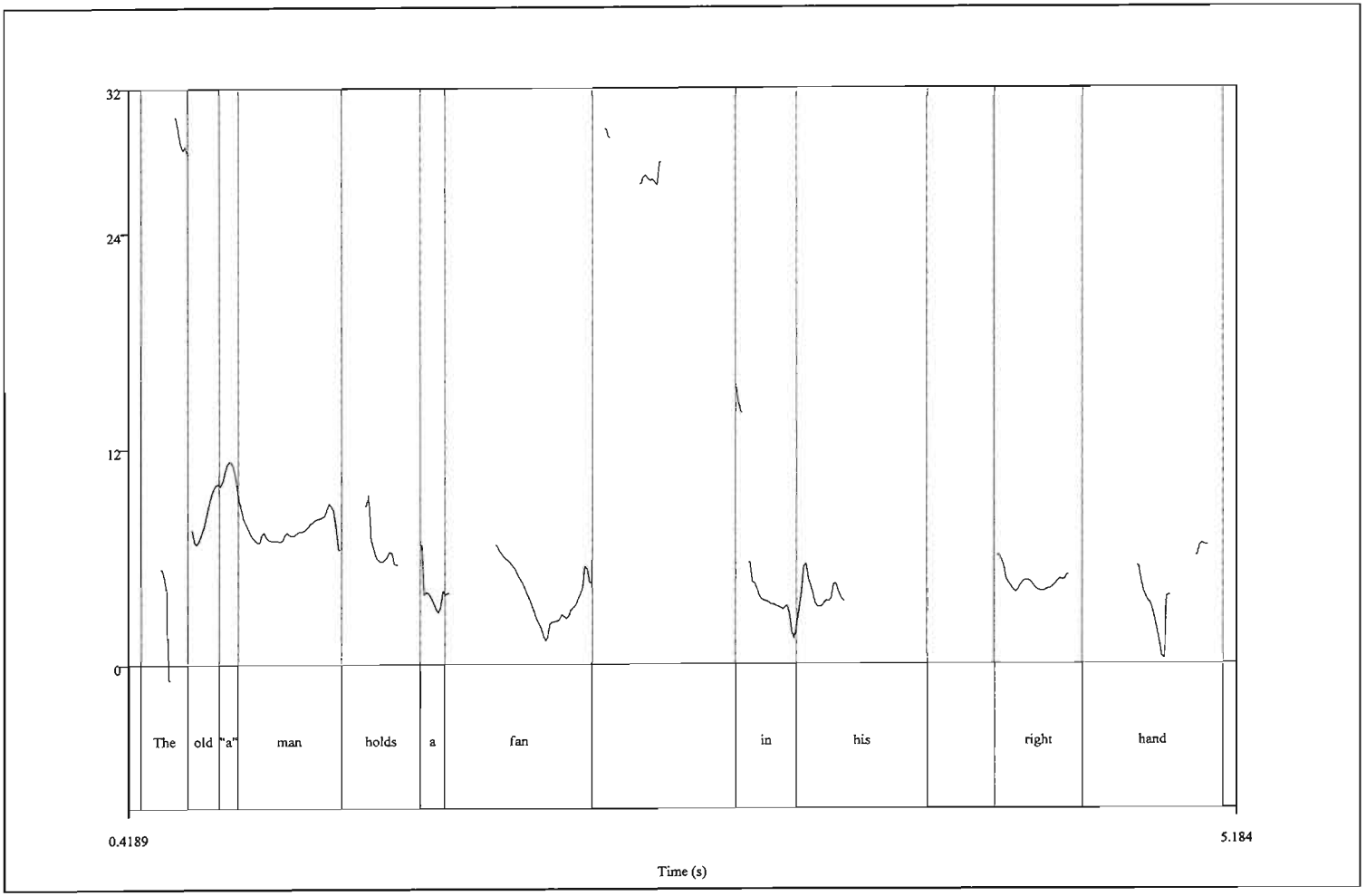
Speaker SC18



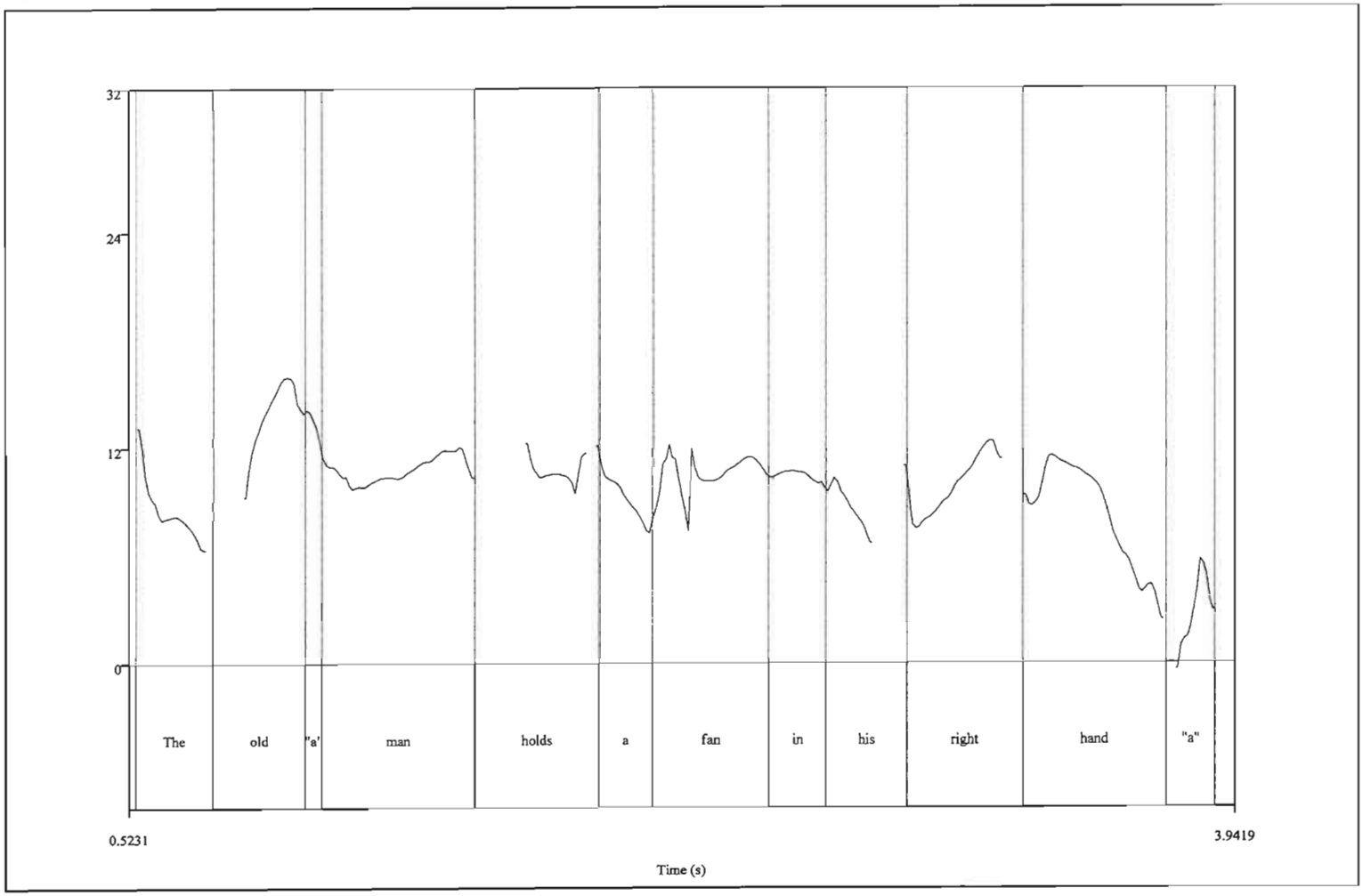
Speaker SC13



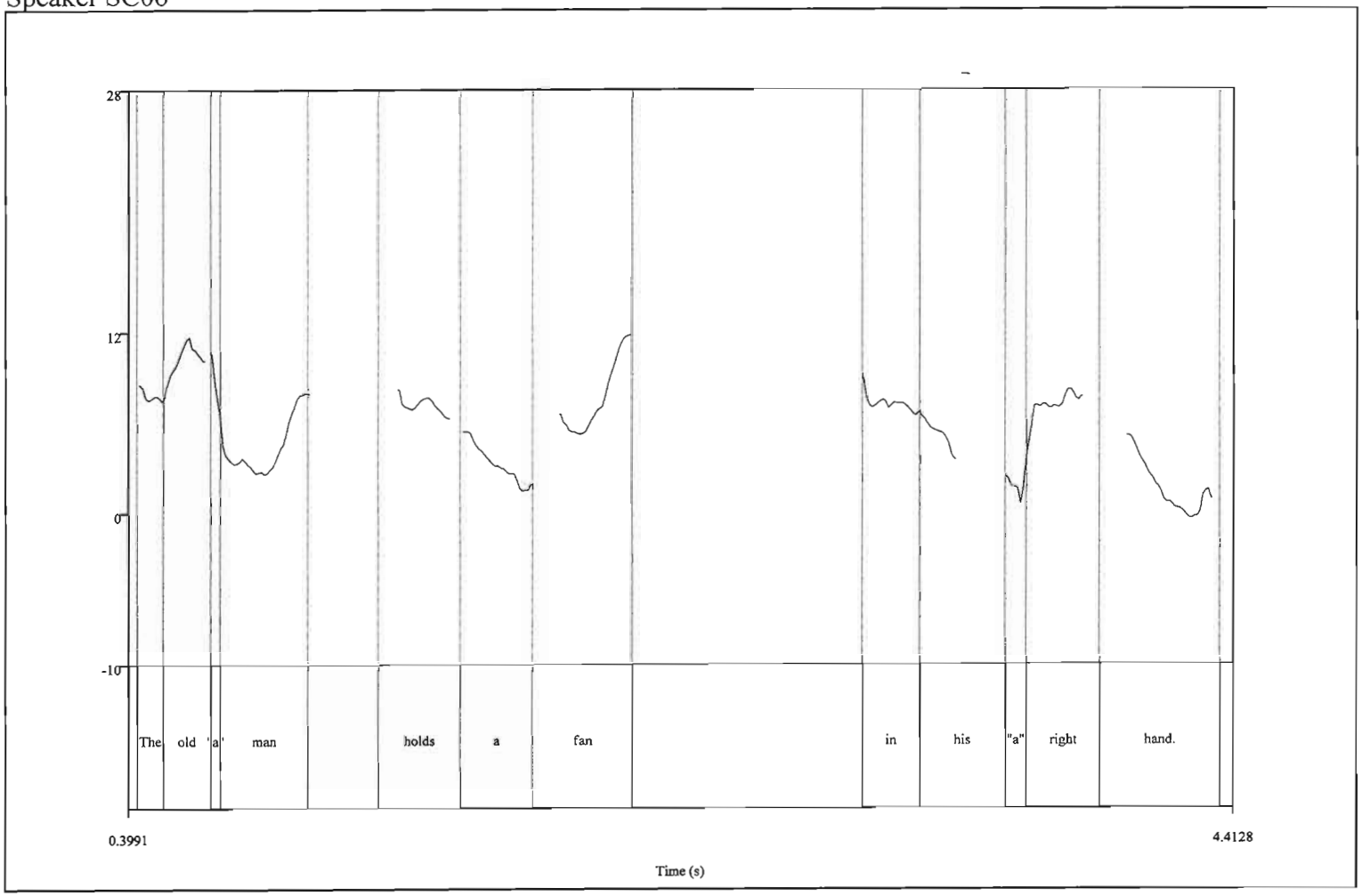
Speaker SC19



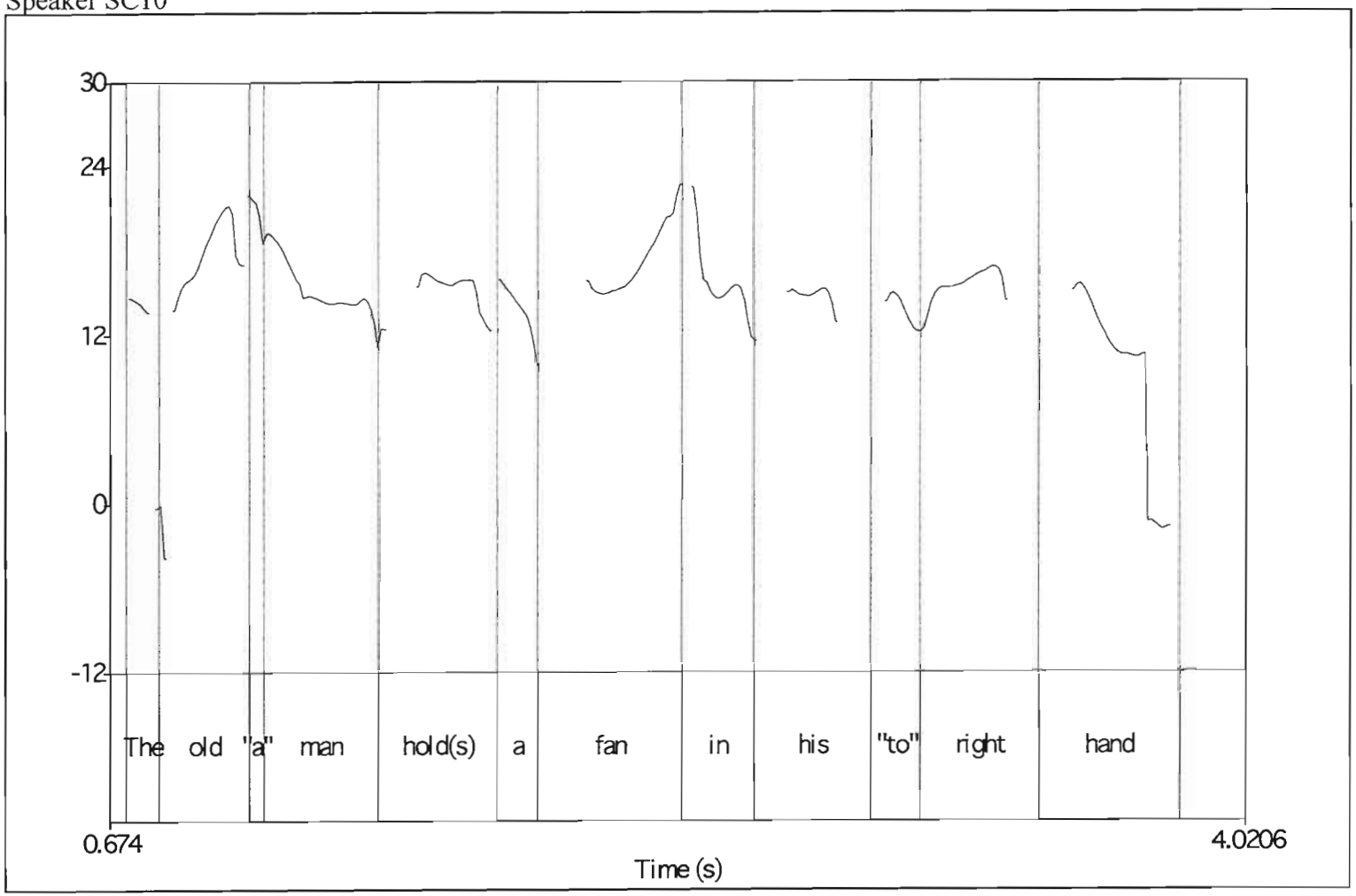
Speaker SC20



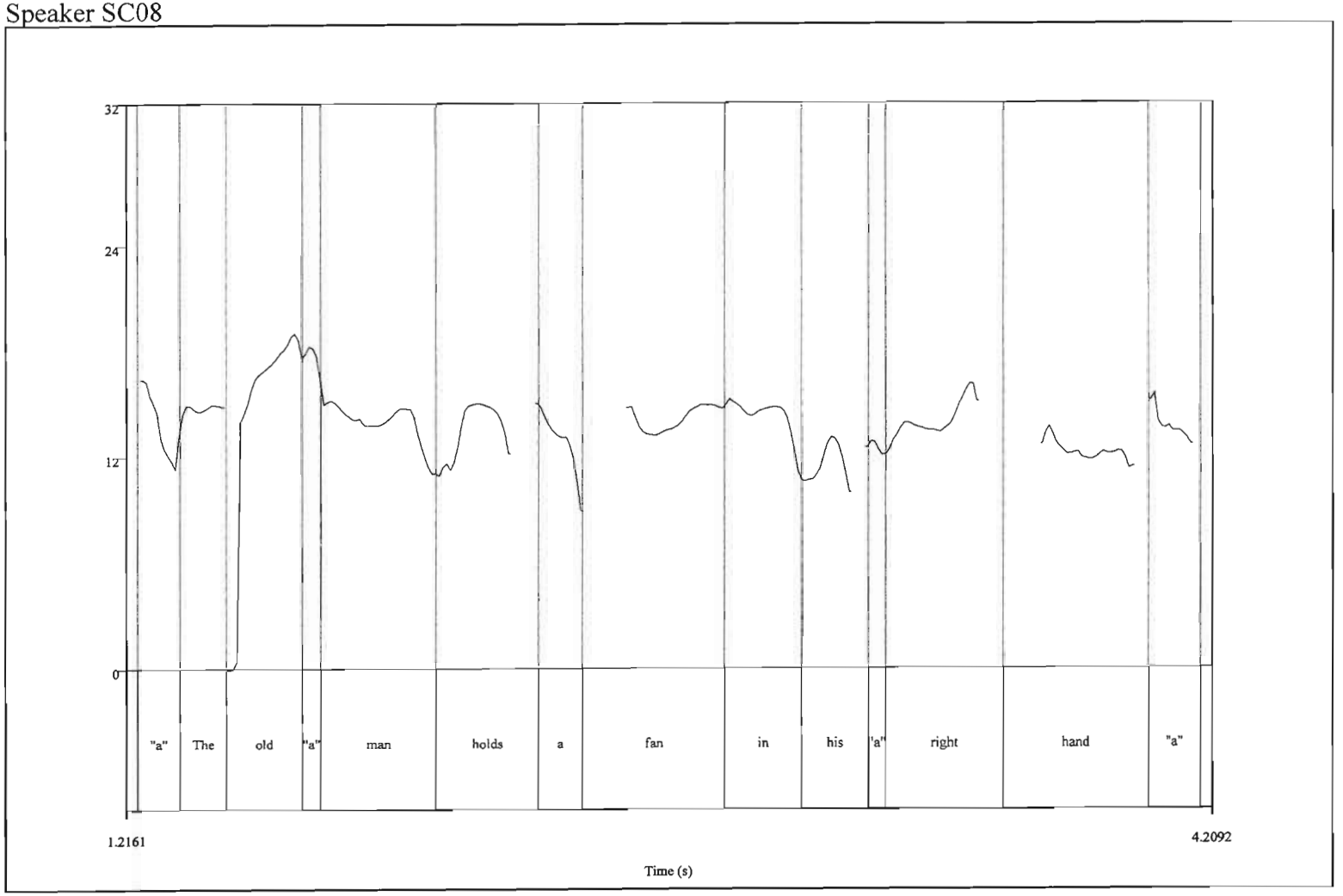
Speaker SC06



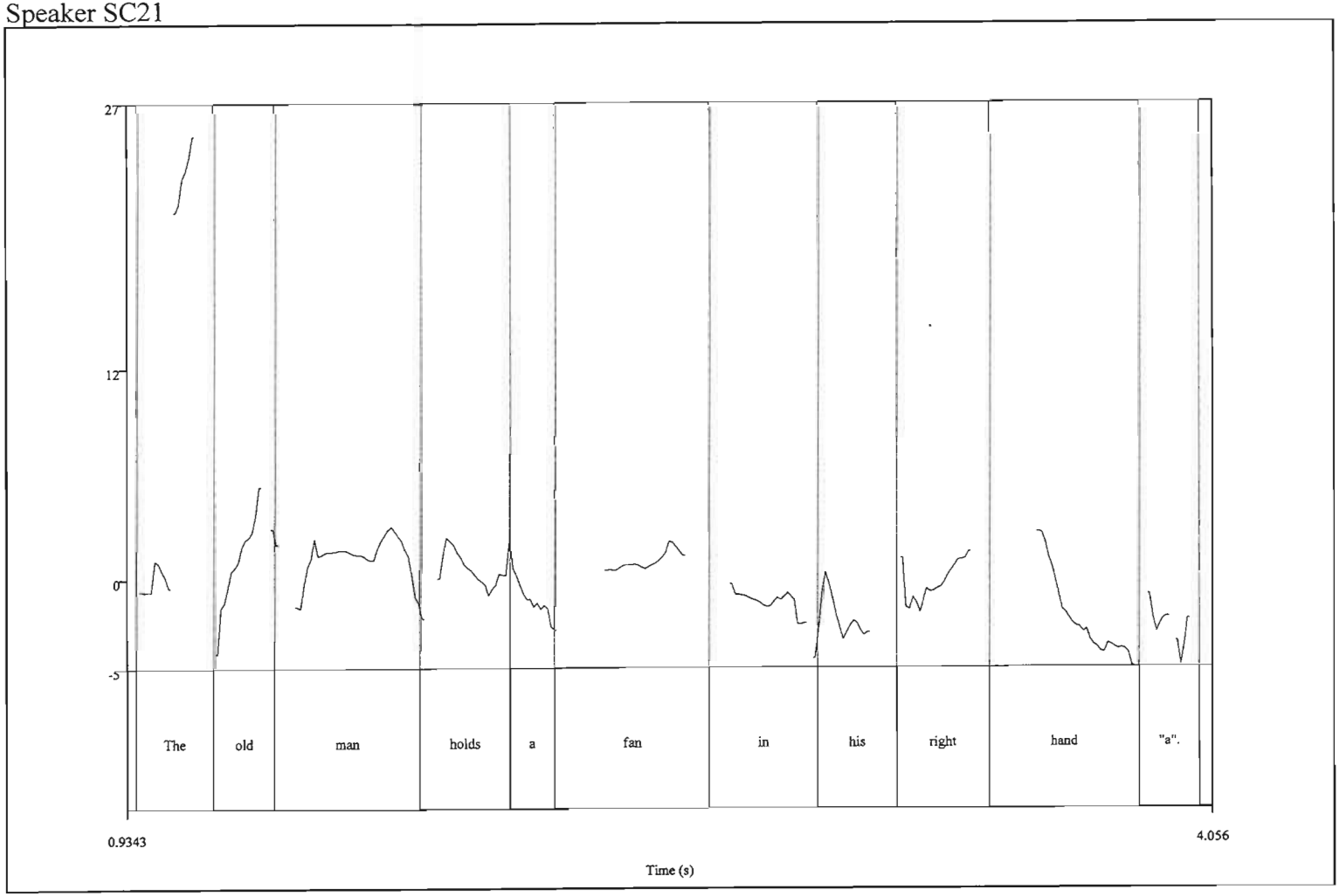
Speaker SC10



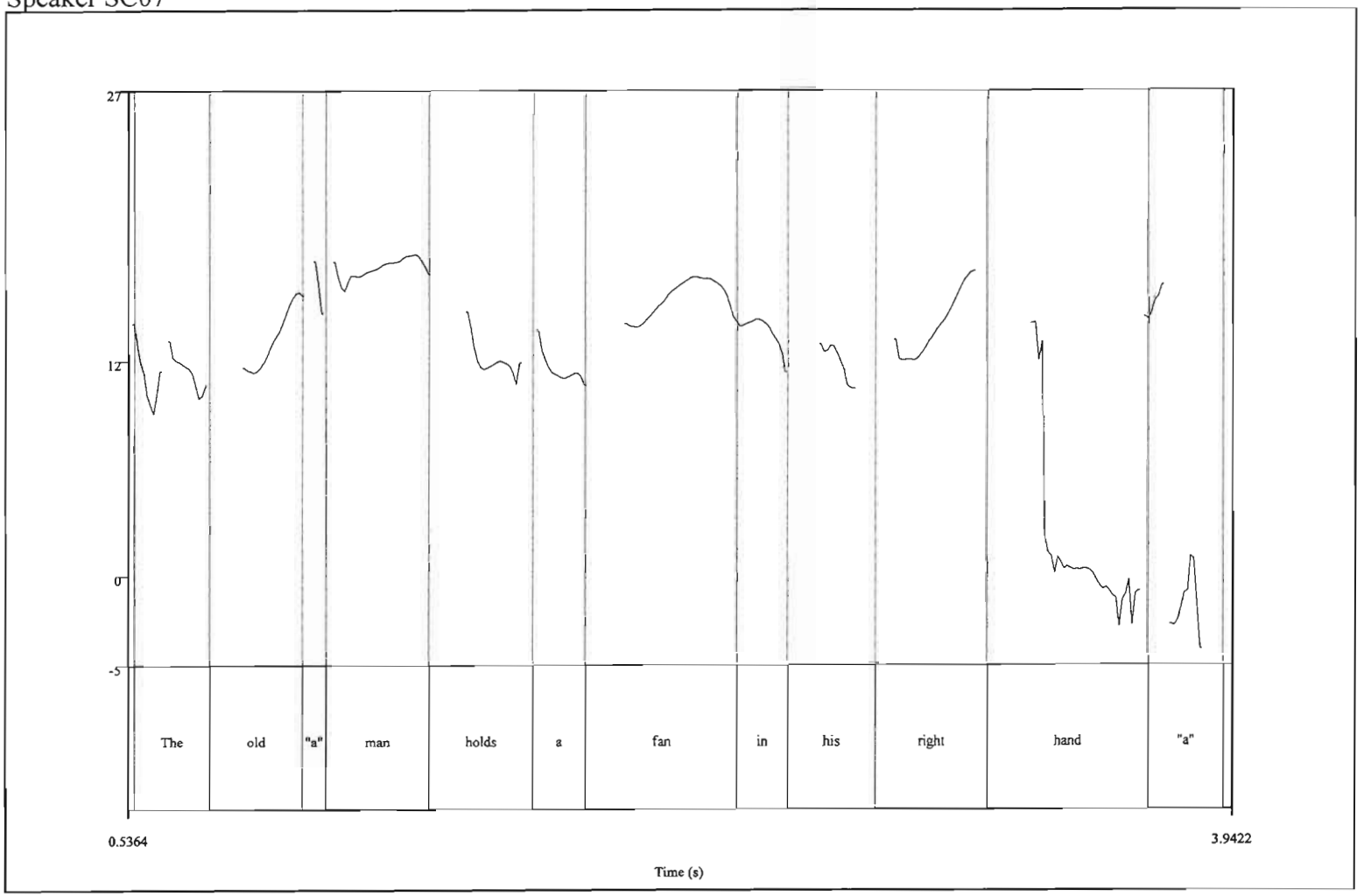
Speaker SC08



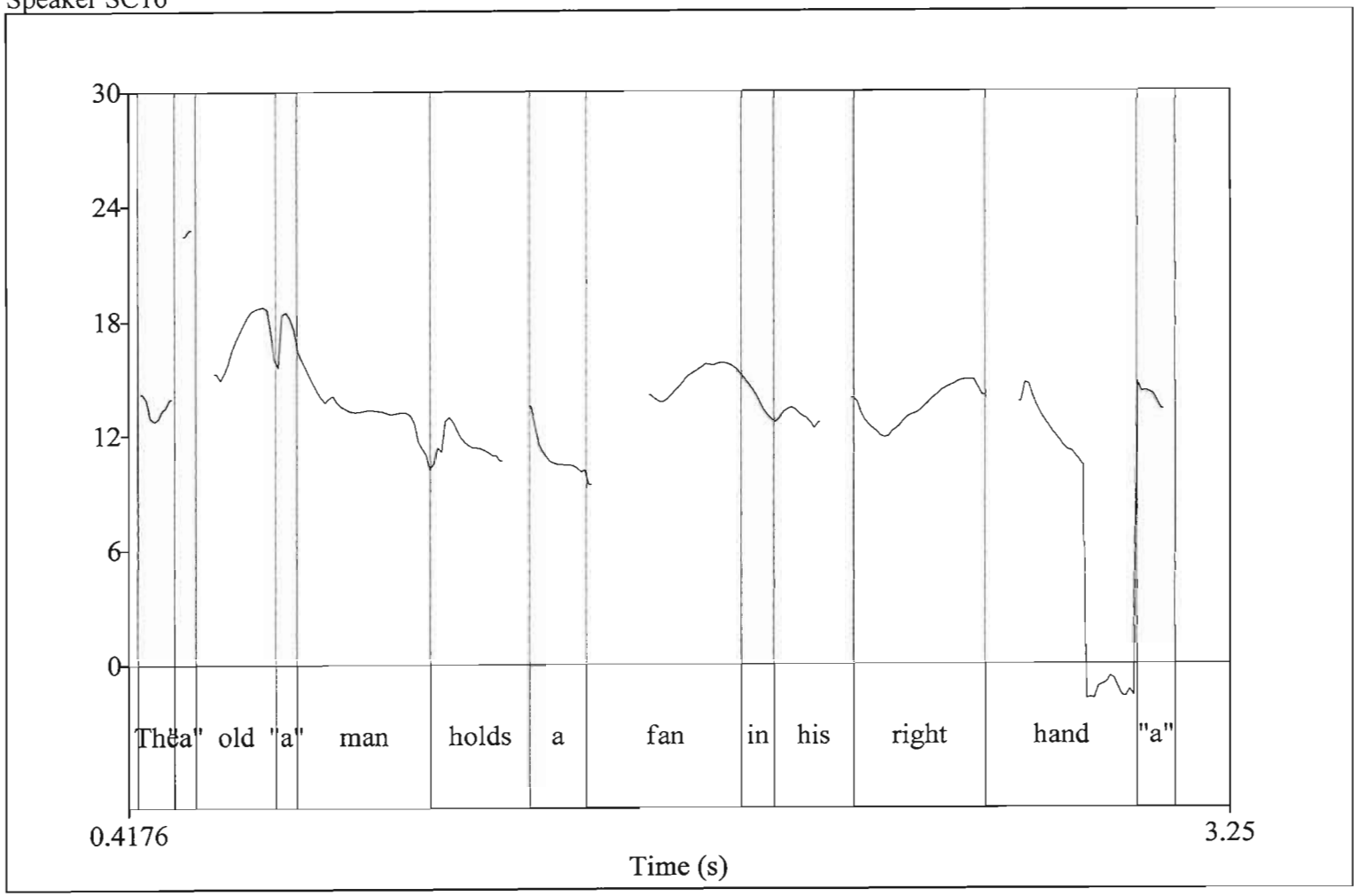
Speaker SC21



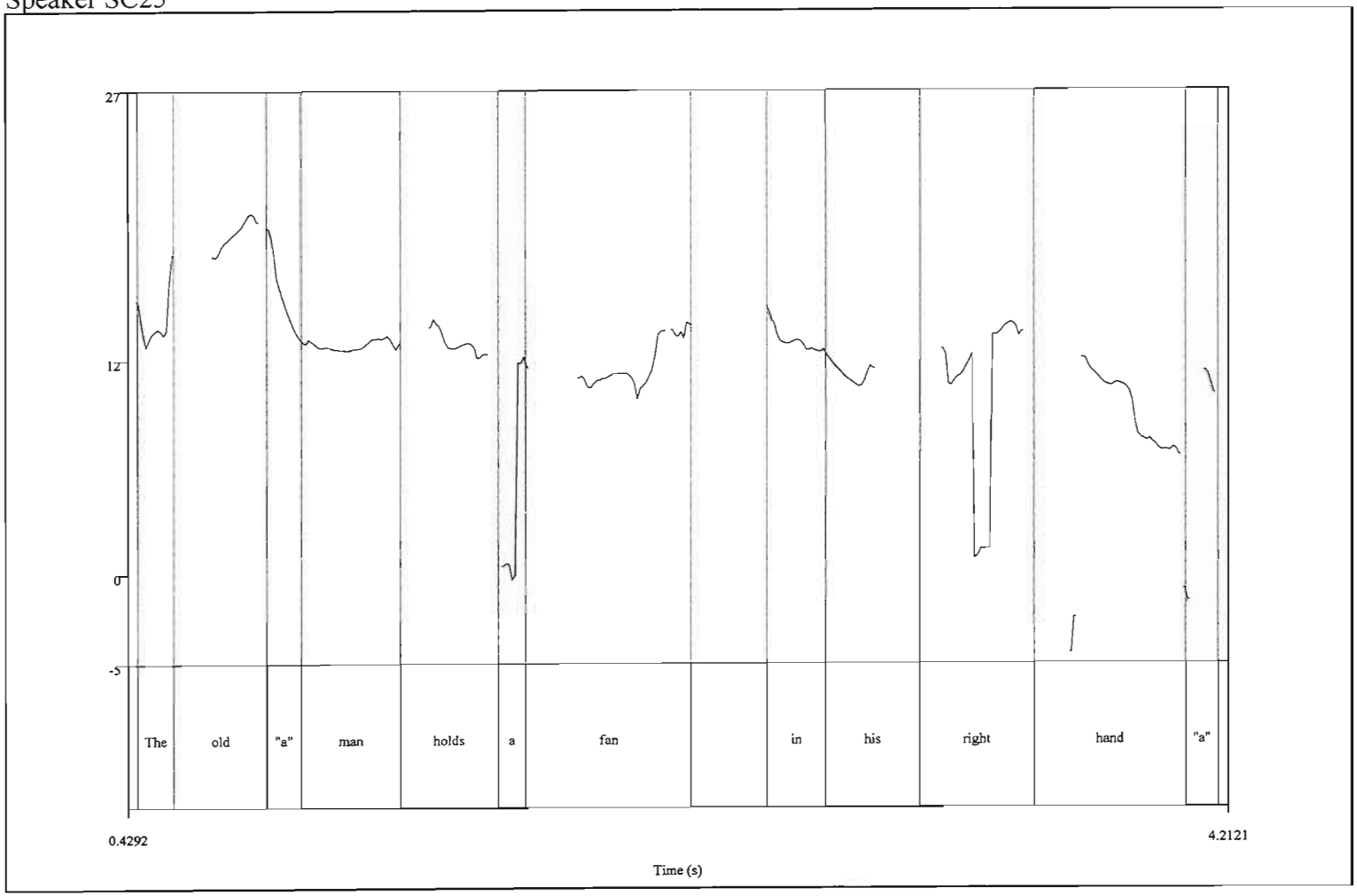
Speaker SC07



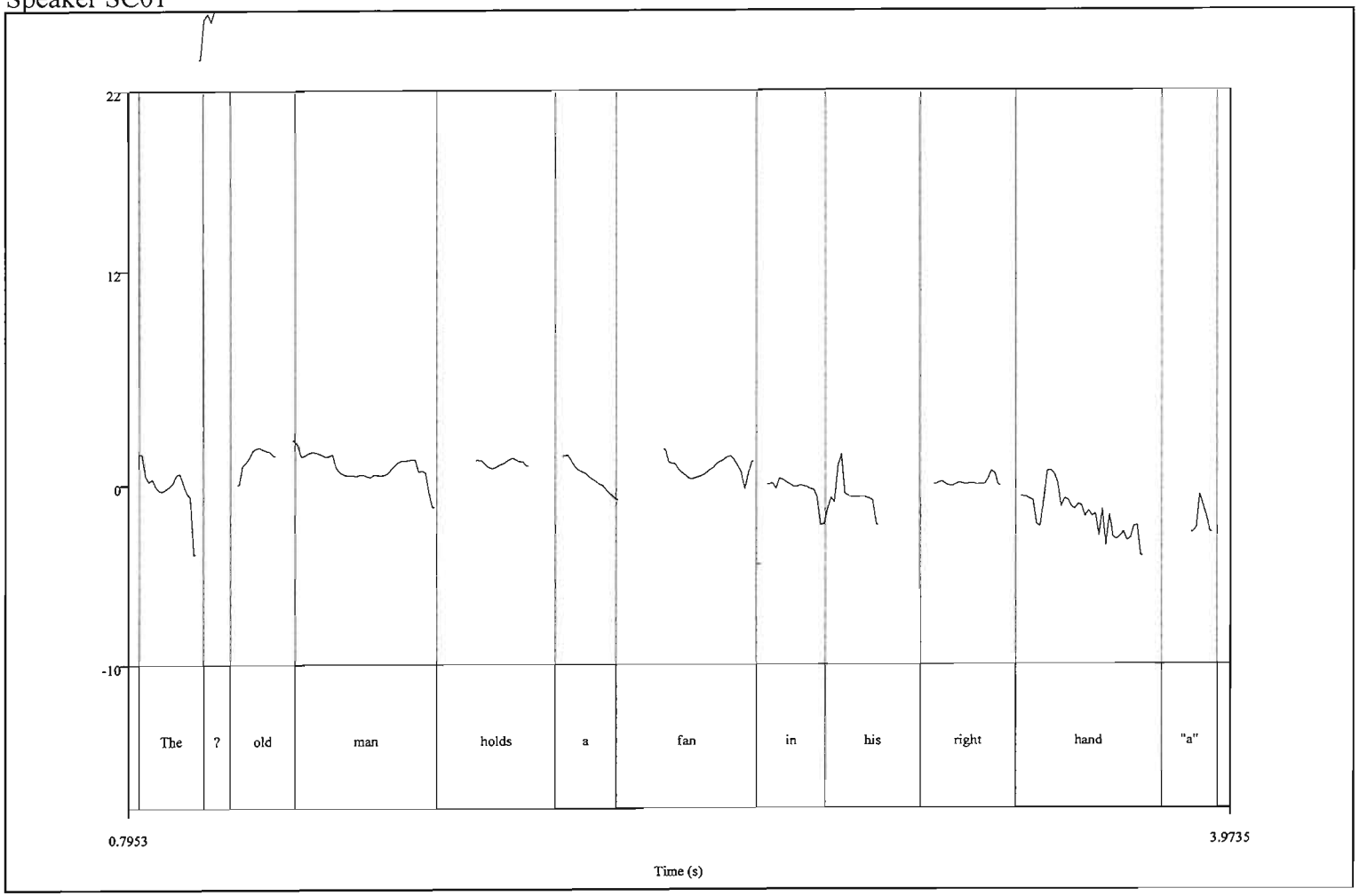
Speaker SC16



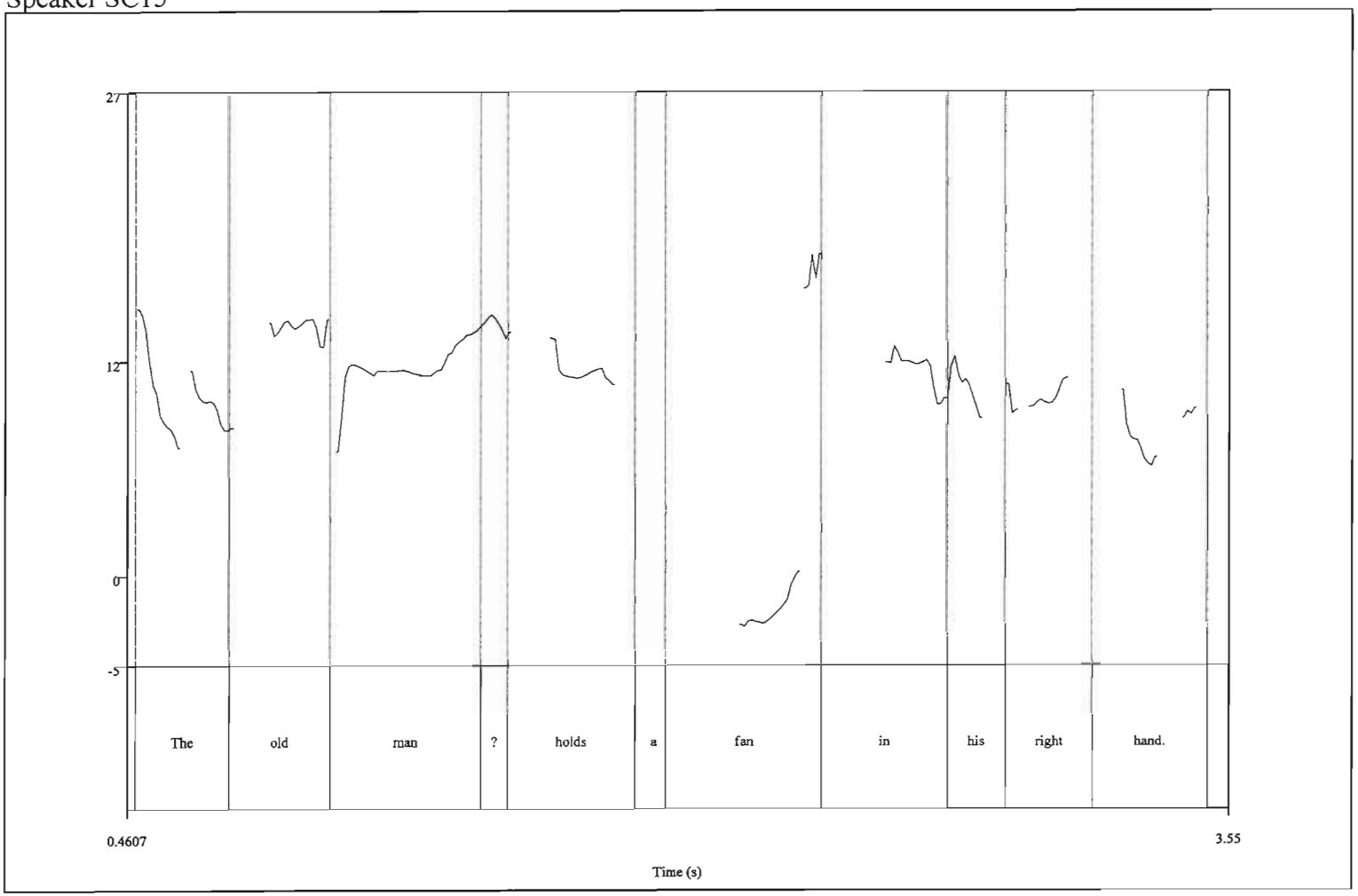
Speaker SC25



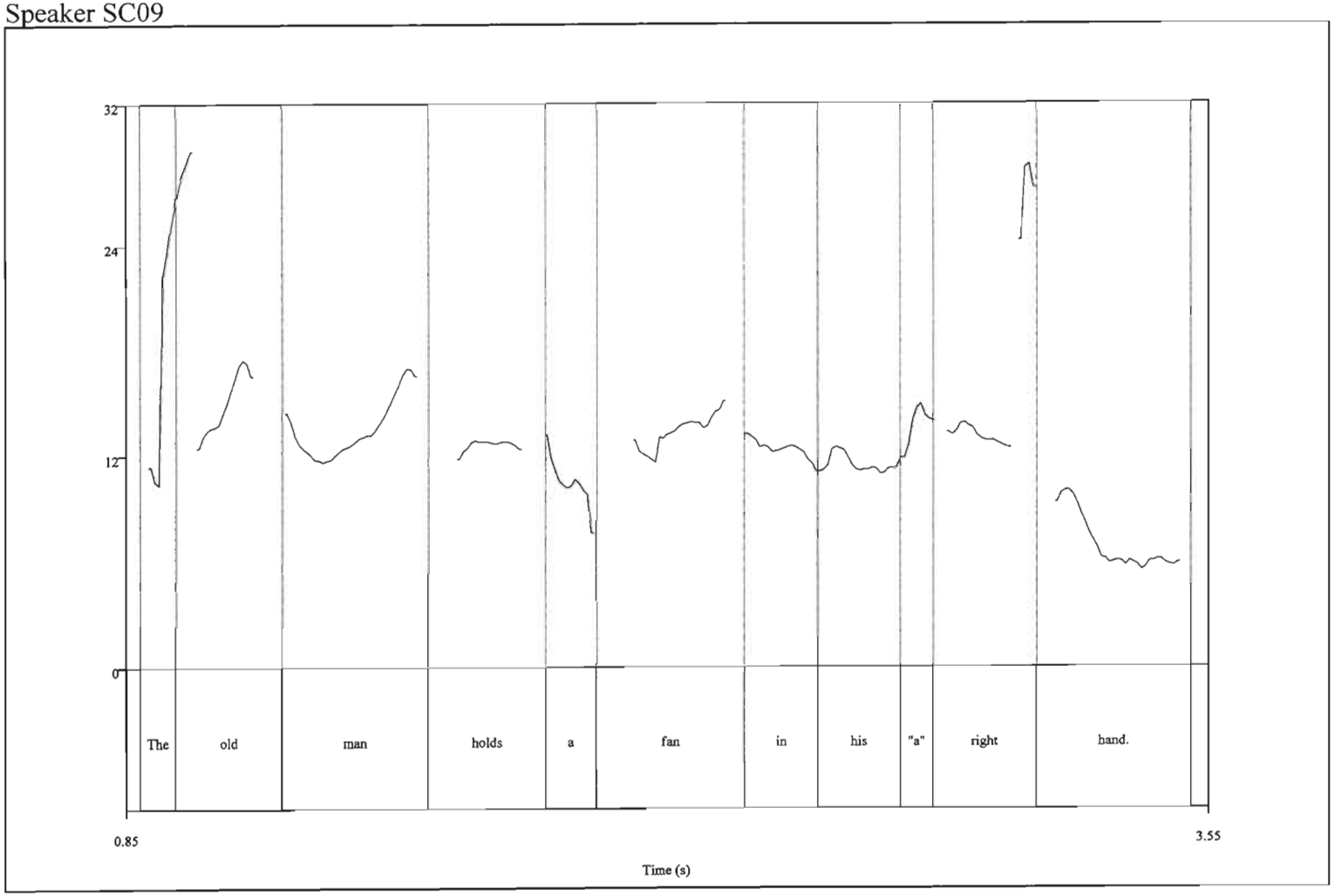
Speaker SC01



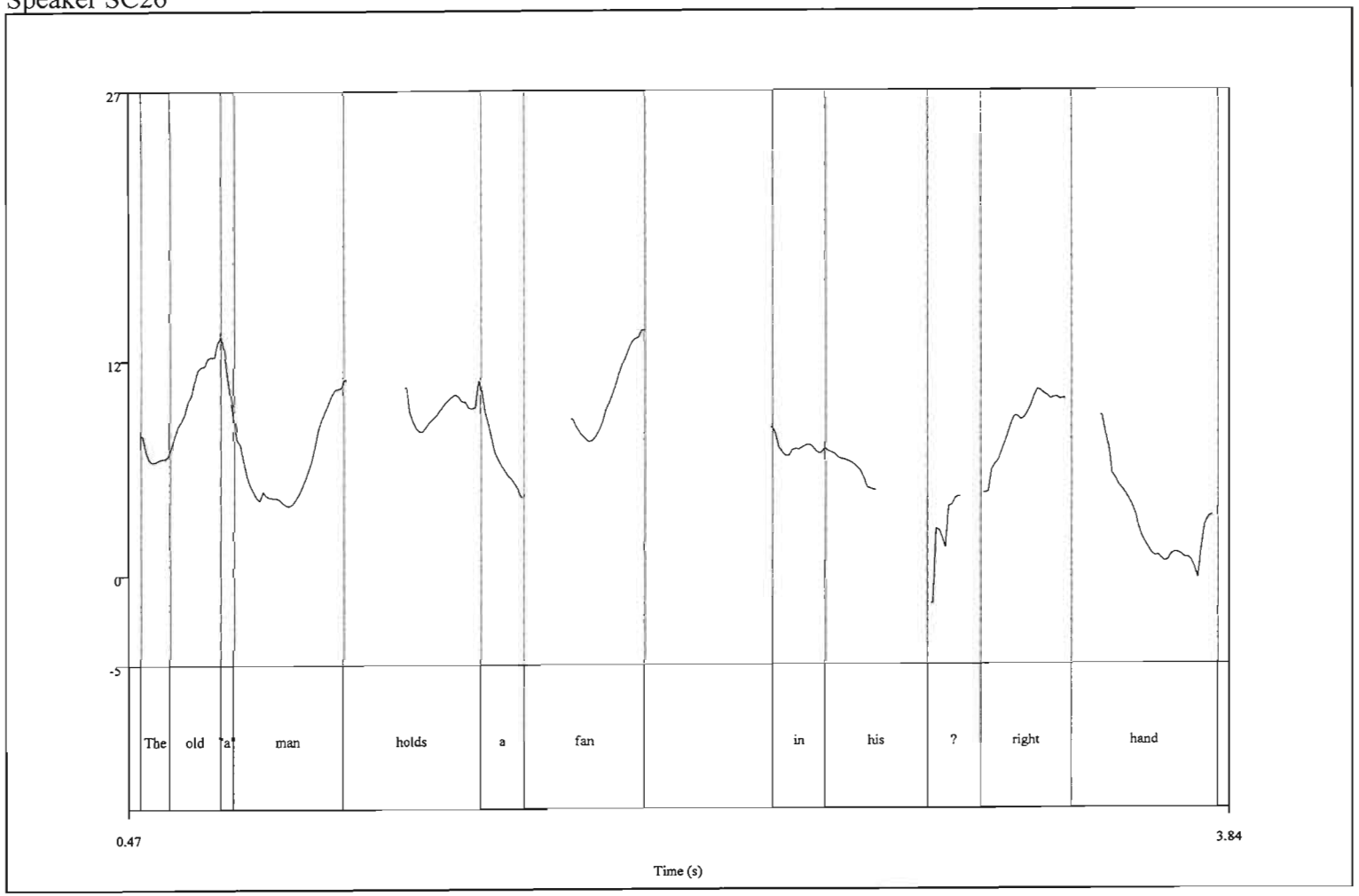
Speaker SC15



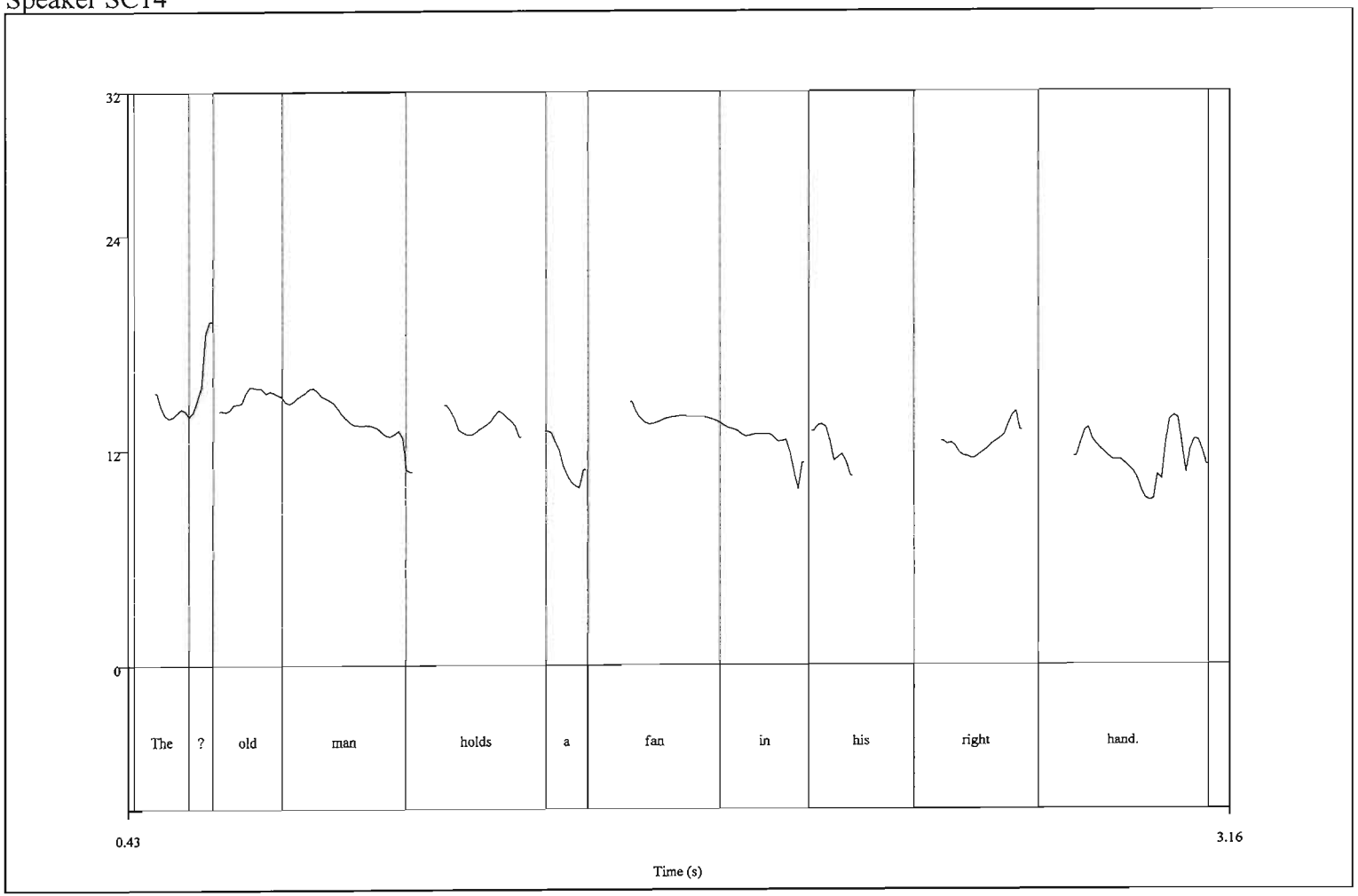
Speaker SC09



Speaker SC26

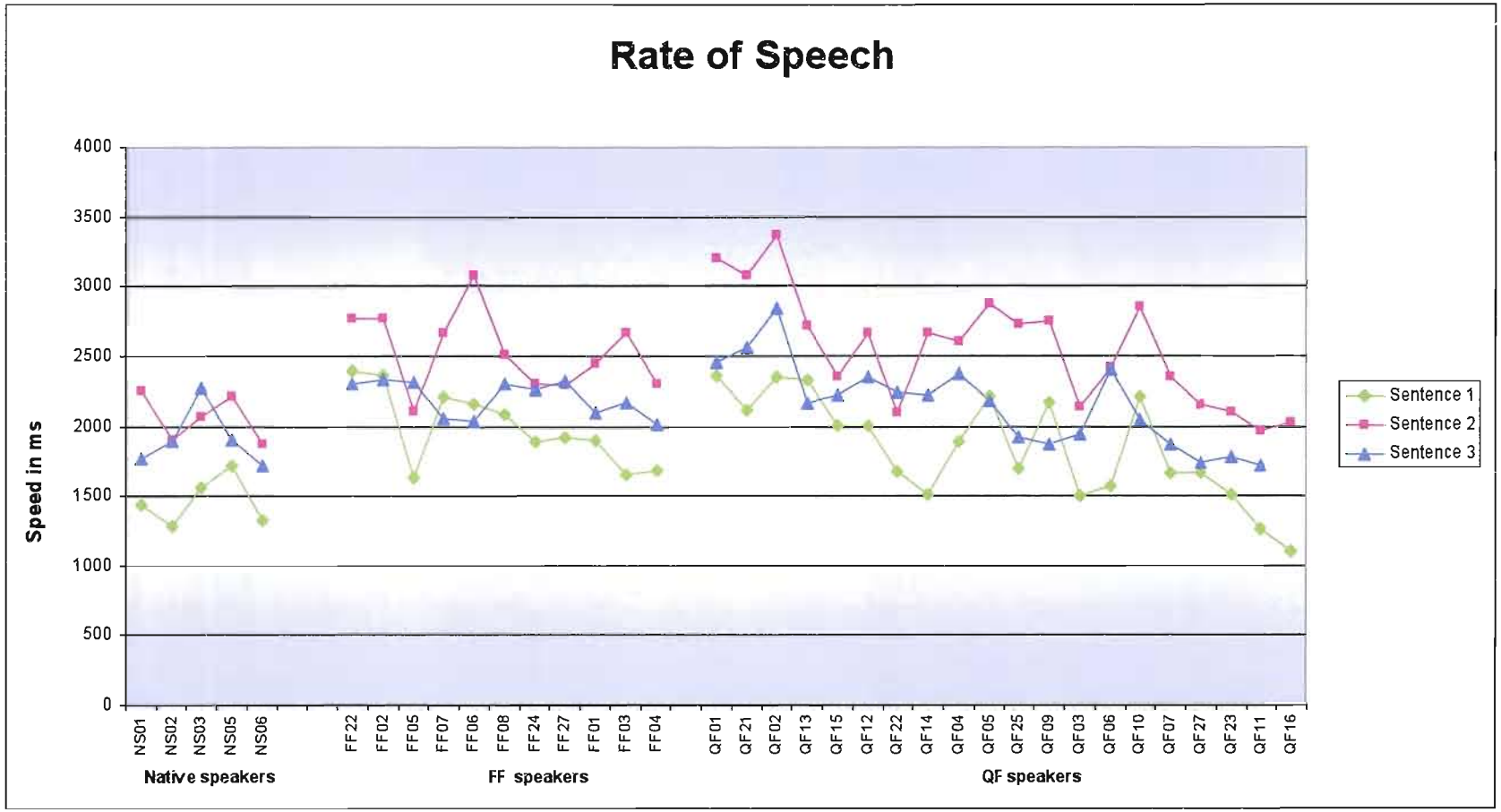


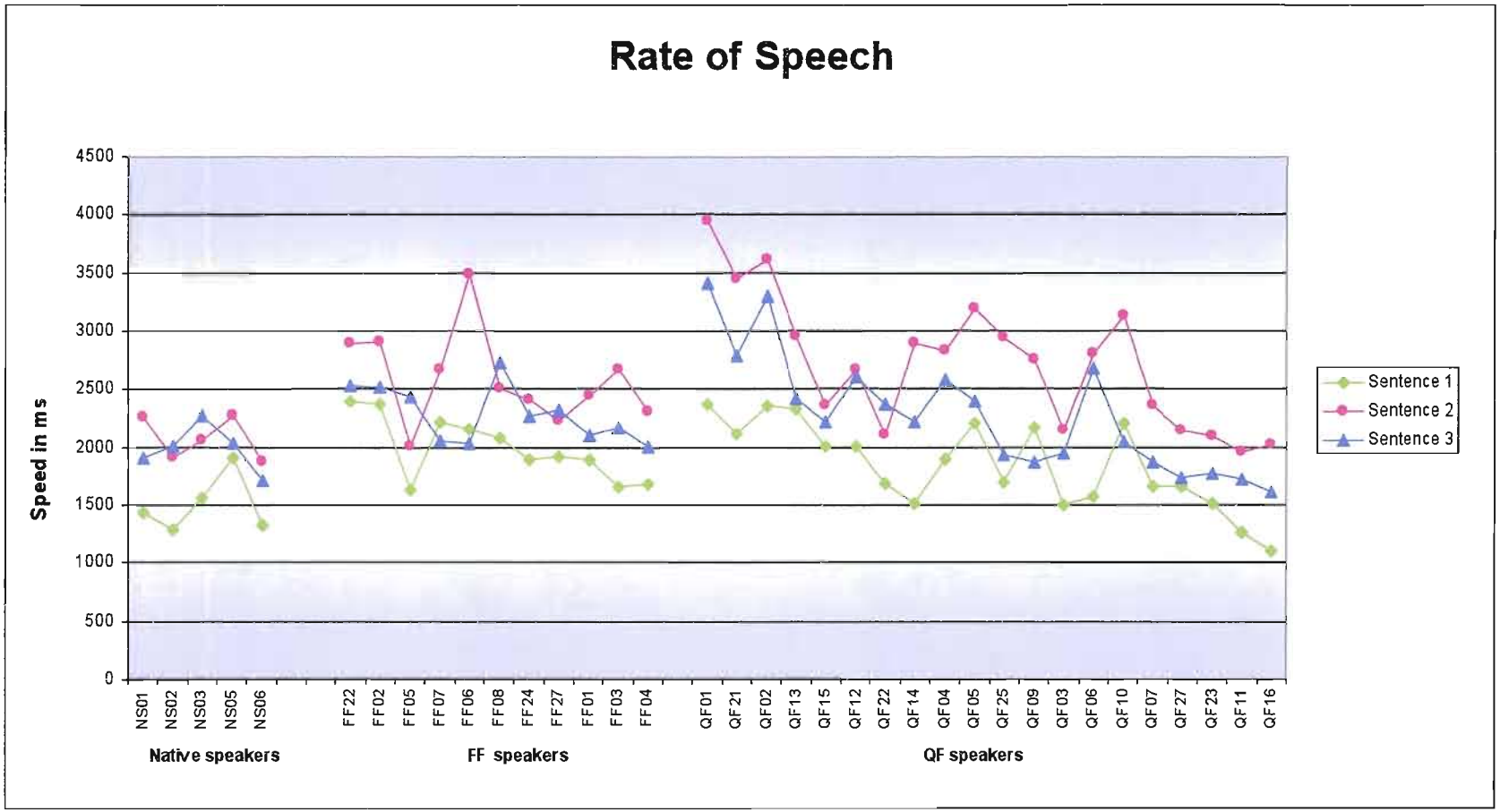
Speaker SC14

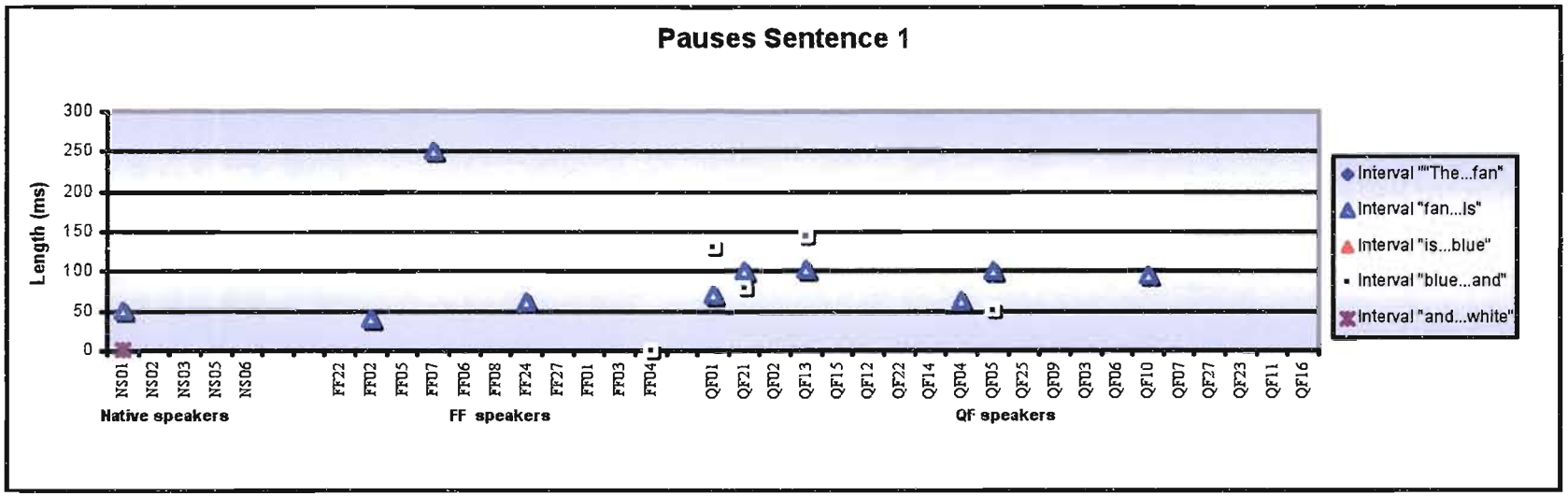


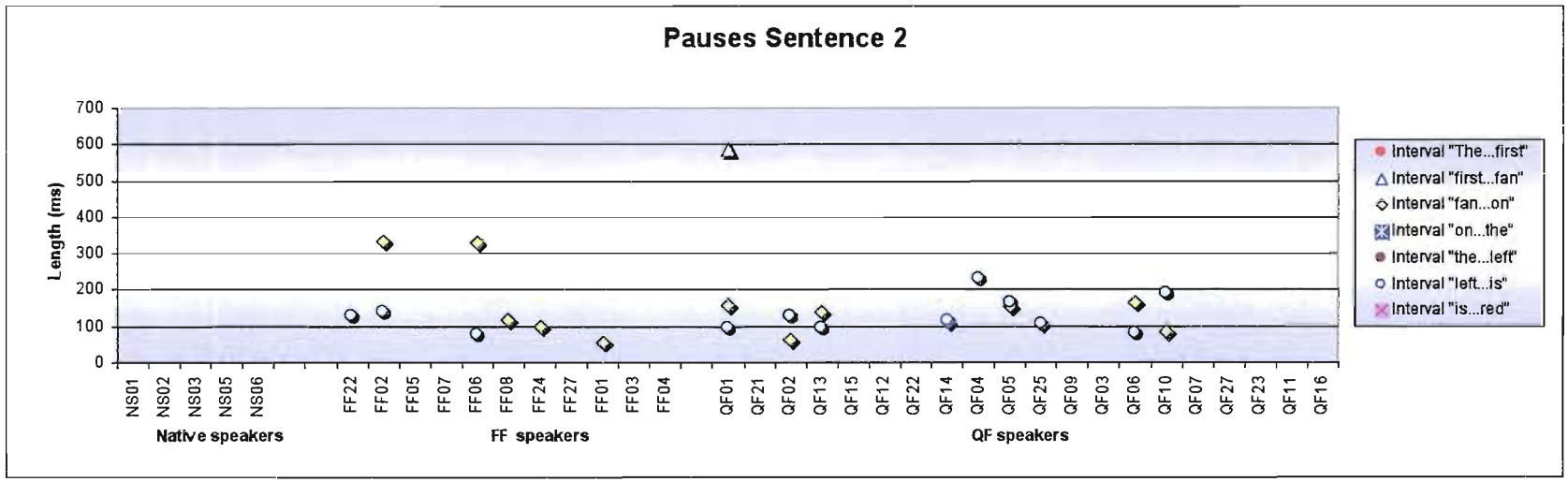
APPENDIX 5

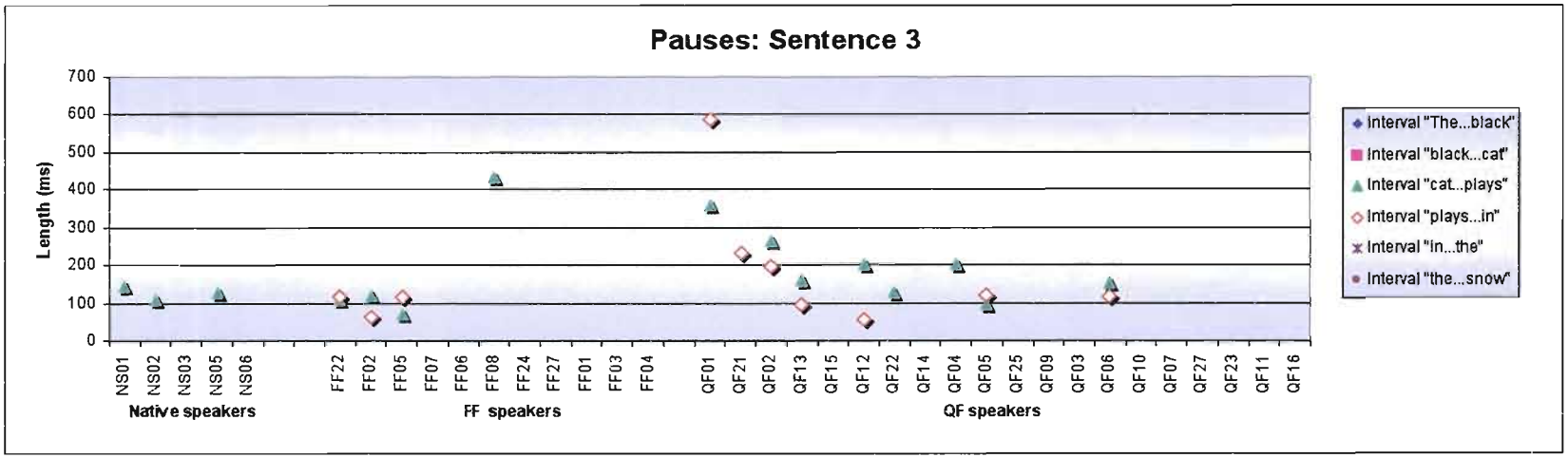
LARGE-SCALE FIGURES

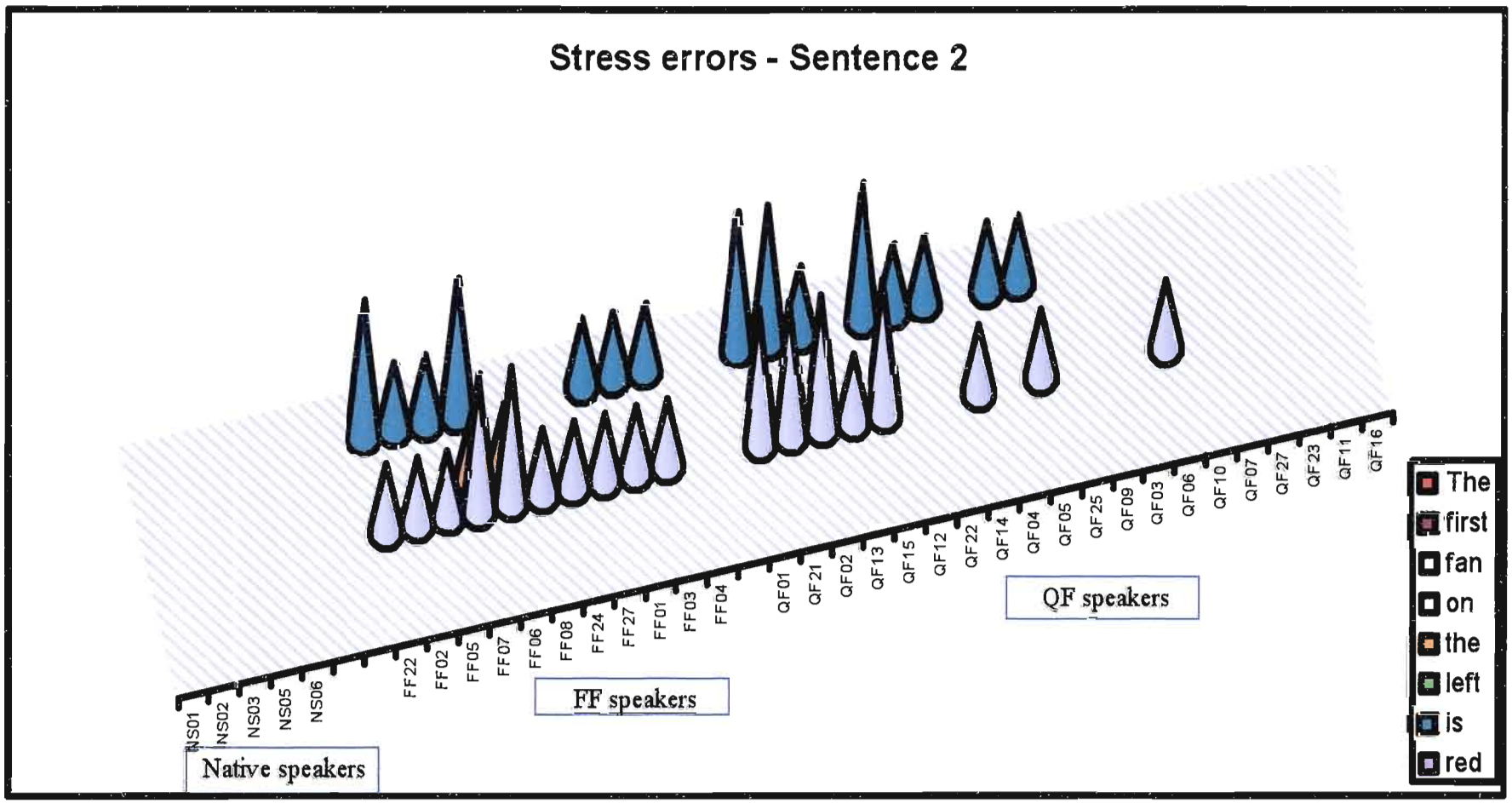

Rate of Speech
Speed in ms
4000
3500
3000
2500
2000
1500
1000
500
0
NS01
NS02
NS03
NS05
NS06
Native speakers
FF22
FF02
FF05
FF07
FF06
FF08
FF24
FF27
FF01
FF03
FF04
FF speakers
QF01
QF21
QF02
QF13
QF15
QF12
QF22
QF14
QF04
QF05
QF25
QF09
QF03
QF06
QF10
QF07
QF27
QF23
QF11
QF16
QF speakers
Sentence 1
Sentence 2
Sentence 3

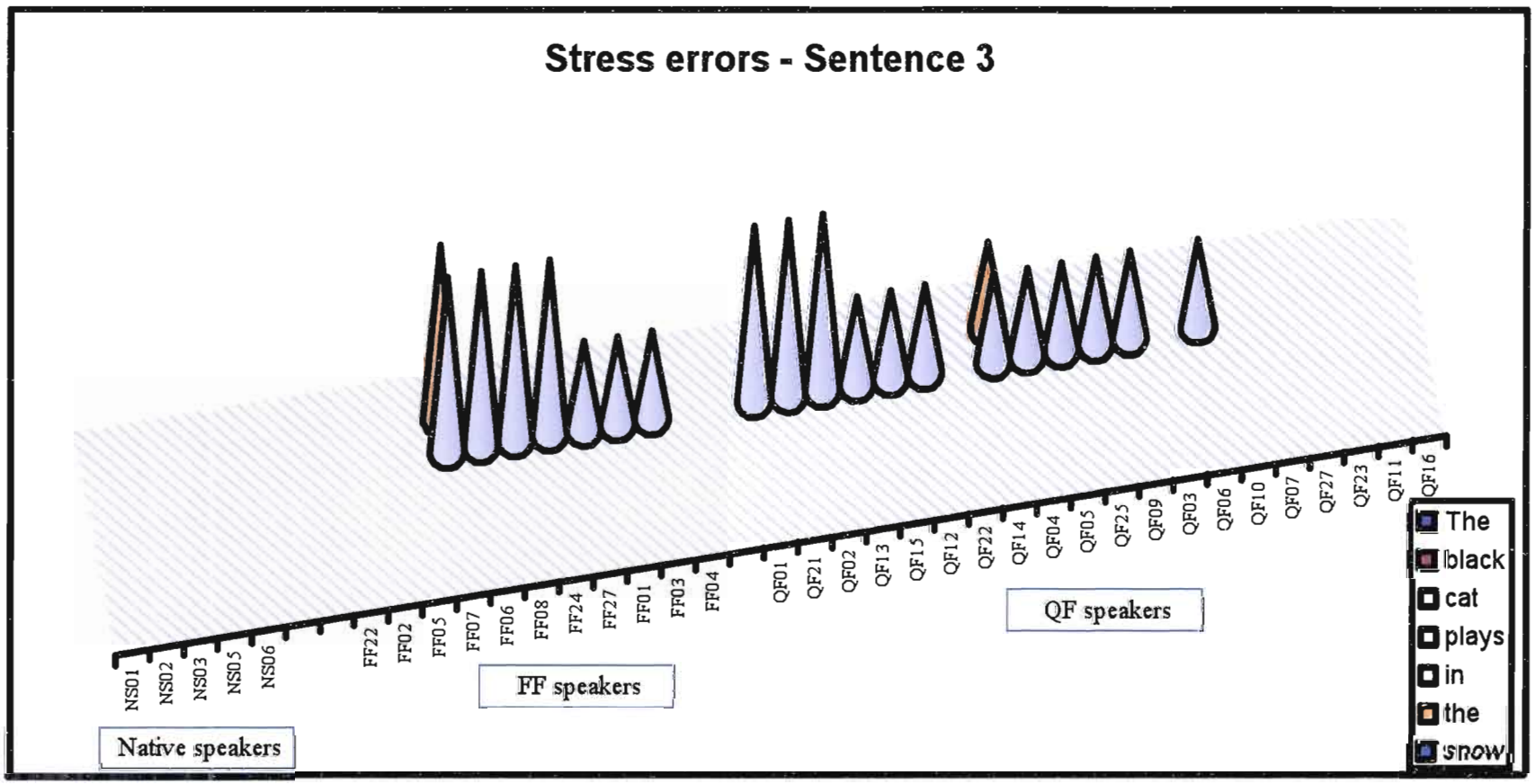

Rate of Speech
Speed in ms
4500
4000
3500
3000
2500
2000
1500
1000
500
0
NS01
NS02
NS03
NS05
NS06
FF22
FF02
FF05
FF07
FF06
FF08
FF24
FF27
FF01
FF03
FF04
QF01
QF21
QF02
QF13
QF15
QF12
QF22
QF14
QF04
QF05
QF25
QF09
QF03
QF06
QF10
QF07
QF27
QF23
QF11
QF16
Native speakers
FF speakers
QF speakers
Sentence 1
Sentence 2
Sentence 3


Pauses Sentence 1
Length (ms)
300
250
200
150
100
50
0
NS01 NS02 NS03 NS05 NS06 FF22 FF02 FF05 FF07 FF06 FF08 FF24 FF27 FF01 FF03 FF04 QF01 QF21 QF02 QF13 QF15 QF12 QF22 QF14 QF04 QF05 QF25 QF09 QF03 QF06 QF10 QF07 QF27 QF23 QF11 QF16
Native speakers
FF speakers
QF speakers
Interval ""The...fan"
Interval "fan...is"
Interval "is...blue"
Interval "blue...and"
Interval "and...white"


Pauses Sentence 2
Length (ms)
700
600
500
400
300
200
100
0
NS01 NS02 NS03 NS05 NS06 FF22 FF02 FF05 FF07 FF06 FF08 FF24 FF27 FF01 FF03 FF04 QF01 QF21 QF02 QF13 QF15 QF12 QF22 QF14 QF04 QF05 QF25 QF09 QF03 QF06 QF10 QF07 QF27 QF23 QF11 QF16
Native speakers
FF speakers
QF speakers
Interval "The...first"
Interval "first...fan"
Interval "fan...on"
Interval "on...the"
Interval "the...left"
Interval "left...is"
Interval "is...red"


Pauses: Sentence 3
Length (ms)
700
600
500
400
300
200
100
0
NS01 NS02 NS03 NS05 NS06
Native speakers
FF22 FF02 FF05 FF07 FF06 FF08 FF24 FF27 FF01 FF03 FF04
FF speakers
QF01 QF21 QF02 QF13 QF15 QF12 QF22 QF14 QF04 QF05 QF25 QF09 QF03 QF06 QF10 QF07 QF27 QF23 QF11 QF16
QF speakers
Interval "The...black"
Interval "black...cat"
Interval "cat...plays"
Interval "plays...in"
Interval "in...the"
Interval "the...snow"


Stress errors - Sentence 2
NS01
NS02
NS03
NS05
NS06
FF22
FF02
FF05
FF07
FF06
FF08
FF24
FF27
FF01
FF03
FF04
QF01
QF21
QF02
QF13
QF15
QF12
QF22
QF14
QF04
QF05
QF25
QF09
QF03
QF06
QF10
QF07
QF27
QF23
QF11
QF16
Native speakers
FF speakers
QF speakers
The
first
fan
on
the
left
is
red


Stress errors - Sentence 3
NS01
NS02
NS03
NS05
NS06
FF22
FF02
FF05
FF07
FF06
FF08
FF24
FF27
FF01
FF03
FF04
QF01
QF21
QF02
QF13
QF15
QF12
QF22
QF14
QF04
QF05
QF25
QF09
QF03
QF06
QF10
QF07
QF27
QF23
QF11
QF16
Native speakers
FF speakers
QF speakers
The
black
cat
plays
in
the
snow

APPENDIX 6

FF SPEAKERS						
Code	**Total points attributed (for rank)** [1]	**Age group**	**Other languages**	**Past training in phonetics**	**Length of study of English**	**Place of study of English**
FF22	14	36-40	Spanish	Yes	0	
FF02	15	36-40	Spanish		0	
FF05	16	26-30			6	France
FF07	17	26-30			1	Quebec
FF06	19	20-25	Arabic		0	
FF08	20	31-35	Arabic	Yes	1	Quebec
FF24	23	20-25	German	Yes	8	France
FF27	24	26-30		Yes	1	Quebec
FF01	26	36-40	Arabic	Yes	0	
FF03	27	31-35	Arabic		4	Morocco
FF04	28	20-25	German		8	France

QF SPEAKERS - IN ORDER OF RANK						
Code	**Total points attributed (for rank)** [1]	**Age group**	**Other languages**	**Past training in phonetics**	**Length of study of English**	**Place of study of English**
QF01	9	20-25				
QF21	13	20-25				Quebec
QF02	15	36-40		Yes	2 months	Quebec
QF13	15	26-30			10	Quebec
QF15	21	26-30			10	Quebec
QF28	21	31-35		Yes	0	Quebec
QF12	22	41-45			5	Quebec
QF22	22	36-40	Spanish	Yes		Quebec
QF14	23	41-45				Quebec
QF04	24	20-25	Spanish		3	Quebec
QF05	24	41-45		Yes	1	Quebec
QF25	24	41-45		Yes	7	Quebec
QF09	29	31-35			4	Quebec
QF03	31	26-30		Yes	10	Quebec
QF06	31	20-25			0	
QF10	32	50-60				Quebec
QF07	34	31-35		Yes	3	
QF27	37	20-25		Yes		Quebec
QF23	38	26-30		Yes		Quebec
QF11	43	26-30	Spanish	Yes	10	Quebec
QF16	44	50-60			10	Quebec

SC SPEAKERS - IN ORDER OF RANK						
Code	**Total points attributed (for rank)** [1]	**Age group**	**Other languages** [2]	**Past training in phonetics**	**Length of study of English**	**Place of study of English**
SC23	8	31-35			10	China
SC04	10	31-35		Yes	2	China
SC11	13	36-40			10	China
SC03	15	31-35	French		7	China
SC12	15	36-40		Yes	1	Quebec
SC02	16	41-45	French		10	China
SC17	17	31-35			3	Quebec
SC18	18	31-35			6	China
SC13	19	41-45			10	China
SC19	20	31-35			8	China
SC20	20	41-45		Yes	10	China
SC06	21	31-35	French	Yes	21	China
SC10	21	26-30	French		10	China
SC08	22	41-45	French		6	China
SC21	22	20-25			6	China
SC07	23	31-35			8	China
SC16	23	31-35			6	China
SC25	24	36-40			6	China
SC01	25	20-25	French	Yes	3	China
SC05	25	36-40	French	Yes	6	China
SC15	25	36-40			10	China
SC09	26	31-35		Yes	16	China
SC26	26	31-35		Yes	21	China
SC14	31	31-35	French		10	China

[1] Minimum number of points: 6; maximum number of points: 36